

FINDPATTERNS on pir: * allowing 0 mismatches

1 <(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,H,E,T)

ORX > end of seq; X = any A/T

Databases searched: NBRF, Release 73.0, Released on 16Aug2002, Formatted on 20Aug2002

Total finds: 0

Total lengths: 96,134,422

Total sequences: 283,224

CPU time: 01:02.36

the Quest program provides the answer sets for the issued, pending & A-GeneSeq files. the reference & alignment data are married to gether in these files.

Finally, there are 2 disks for seq1 run on computer & up the & X Scoring table.

Ex. pattern, the GCG program, find patterns, search both the PIR & Swissprot databases. At first, I searched the sequences w/ no limitations on length of the database sequence. the files were huge & unmanageable, then I searched exact length (see left). there were no answer in SWISSPROT or PIR for the exact length seq.

then I learned the sequence w/ 20 A/T's on either side. there are answer sets for each data & base the seq2supl.find (or seq2pir1.find) file shows the alignments & the accession number of the database seq. the .ref files give the information on the accession number. (Sorry I can't marry them together.)

Point of Contact:

Susan Hanley

Technical Info. Specialist

CM1 6B05 Tel: 305-4053

Find patterns on swp: * allowing 0 mismatches

sequence for all find patterns run

any of these 61 times

any (A,D,E,I) DR X {1,20} end of sequence

```

1 1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
AL20_CARMA ck: 9046 len: 27 i P81823 carinus maenas (common shore crab)
1: GYEDEDERPFYALGGRKPRRTYSFGL
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
AMD1_CHICK ck: 7006 len: 26 i P81073 gallus gallus (chicken). amp deamina
1: MNOKHLRFIKKSIVYDADRVVYDAK
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
AMD1_RABIT ck: 7826 len: 26 i P81072 oryctolagus cuniculus (rabbit). amp
1: MNOKHLRFIKKSIVYDADRVVYDAK
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
DNIV_BPDI0 ck: 9004 len: 22 i Q38199 bacteriophage d108. dna-invertase (f
1: YKHPAKRTHIENDRINOIDR
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
FIBB_ANTAM ck: 7016 len: 21 i P14465 antilocapra americana (pronghorn). f
1: QPSYDYEEDDDRRAKLRLDAR
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
FIBB_BISBO ck: 7213 len: 21 i P14466 bison bonasus (european bison). fibr
1: EFPTDYDEGEDDRPKVIGGAR
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
FIBB_BUBBU ck: 7185 len: 21 i P14467 bubalus bubalis (domestic water buff
1: QFPTDYDEGEDDRPKLIGGAR
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
FIBB_CEREL ck: 6821 len: 21 i P14468 cervus elaphus (red deer), and cervu
1: QHSTDYDEEDDRAKLHLHDAR
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
FIBB_FELCA ck: 5816 len: 20 i P14469 felis silvestris catus (cat). fibrin
1: IIDYDEGEDDRDGVVDAR
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
FIBB_LAMGL ck: 4317 len: 19 i P14473 lama glama (llama), lama vicugna (vi
1: ATDIDEEDDRKYVRLDAR
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
FIBB_MUNMU ck: 6363 len: 21 i P14475 muntiacus muntjak (muntjak). fibr
1: QHSTDYDEEDDRAKLHLHDAR
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
FIBB_ODOHE ck: 6856 len: 21 i P14476 odocoileus hemionus (mule deer) (
1: QHLDYDEVEDDRRAKLHLHDAR
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
FIBB_RANTA ck: 6866 len: 21 i P14479 rangifer tarandus (reindeer) (car
1: QHLDYDEVEDDRRAKLHLHDAR
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
FIBB_SYNCA ck: 7177 len: 21 i P14481 syncerus caffer (cape buffalo). f
1: QFPTDYDEGEDDRPKSGLGAR
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
GLU1_ORENI ck: 382 len: 36 i P81026 oreochromis niloticus (nile tilap
1: HSEGTFSNDYSKYLEDRKAODFVRLMNNKRSAGAE
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
LPGF_ECOLI ck: 4518 len: 19 i P33236 escherichia coli. gef leader pept
1: MLNTCRVPLTDRKVKERKA
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
MY14_EISFO ck: 7600 len: 14 i P46979 eisenia foetida (common brandling
1: GFKDGADDRISHGP
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
MY14_PHEVI ck: 7729 len: 14 i P46980 pheretima vittata (earthworm). my
1: GFRGDSADDRISHGP
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
T2A_PARTE ck: 946 len: 23 i Q27173 paramoecium tetraurelia. trichocys
1: DPIDRLSTLTDLIEDRYVAEQKE
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
T2B_PARTE ck: 1016 len: 23 i Q27174 paramoecium tetraurelia. trichocys
1: DPIDRLVSTLTDLIEDRYVAEQKE
1 <X[0,20](A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}(A,D,E,I)DRX{18}
Q9ZG33 ck: 1658 len: 17 i Q9ZG33 chlamydia trachomatis. subtilisin

```

1	1:	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(8){T}DRx(6) NFLGPFCTDRKNSFRI	1	1:	GSDFASQSLKURLSDRTADTNRIRIKIIMRYLNS
1	1:	Q9NZ25 ck: 6136 len: 20 i Q9nz25 homo sapiens (human). atp7b (frag <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(6){A}DRx(11) APIQQLADRTSGYFPEYIV	1	1:	Q9NZ25 ck: 6136 len: 20 i Q9nz25 homo sapiens (human). atp7b (frag <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(6){A}DRx(11) APIQQLADRTSGYFPEYIV
1	1:	Q9KI16 ck: 2042 len: 24 i Q9ki16 streptomyces coelicolor a3(2). fblr <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(A){A}DRx(19) MAADRTGTEAATAERALGSRAPEF	1	1:	Q15244 ck: 2868 len: 40 i Q15244 homo sapiens (human). phosphoribo <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(19){A}DRx(18) MPNKKIFSGSSHODLSOKIADRLGLGKVVYTKKFSNQET
1	1:	Q9X9U4 ck: 4039 len: 25 i Q9x9u4 streptomyces coelicolor. adp glucose <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(18){A}DRx(4) MLGIYLAGGEGKRLMPLTADRAKPA	1	1:	Q15245 ck: 3314 len: 40 i Q15245 homo sapiens (human). phosphoribo <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(19){A}DRx(18) MPNIVLFSGSSHODLSQRYADRLGLGKVVYTKKFSNQET
1	1:	Q49137 ck: 9647 len: 32 i Q49137 methyllobacterium extorquens. mxas pr <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(16){A}DRx(13) ALRRICAPFARPPFRLADRFDAELSRHLMWT	1	1:	Q9BQV8 ck: 5960 len: 34 i Q9bqv8 homo sapiens (human). l-isoaspart <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(14){T}DRx(16) NGIITKDTKVEVMLADTDRSHYAKCNPMYDSPOSI
1	1:	Q53914 ck: 3396 len: 9 i Q53914 streptomyces cyaneus (streptomyces c <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N (1){T}DRx(5) ITDRRAAQ	1	1:	Q9UC11 ck: 8961 len: 15 i Q9uc11 homo sapiens (human). tropomyosin <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(2){E}{A}DRx(7) HIAEDADRKYEVAR
1	1:	Q50476 ck: 7152 len: 13 i Q50476 mycobacterium tuberculosis. catalase <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(8){T}DRx(2) MPEQHPPTDRSR	1	1:	Q96F68 ck: 588 len: 28 i Q96f68 homo sapiens (human). unknown (pr <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(10){A}DRx(15) VSQPGSCRHGADRVGHGQACAGAVRPE
1	1:	005422 ck: 6757 len: 38 i 005422 mycobacterium marinum. phytoene dehy <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(15){A}DRx(19) VPGVGPTTLISGRLADRTGNTTRTSIRHLDLKAQLS	1	1:	Q9NBB1 ck: 6662 len: 26 i Q9nbb1 drosophila melanogaster (fruit fl <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(8){T}DRx(15) VVARNNGTDRIFVTYTIKIPKKKEE
1	1:	Q93A12 ck: 9419 len: 22 i Q93a12 thiobacillus ferrooxidans. putative <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(9){D}DRx(10) AQNFINPDLDRTEKDLATSTV	1	1:	Q9GU45 ck: 9325 len: 27 i Q9gu45 sycon raphanus. homeobox protein <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(20){T}DRx(4) KTSKYLAFTTRQALAAARLGLTDROVKV
1	1:	Q93A08 ck: 5753 len: 12 i Q93a08 thiobacillus ferrooxidans. resb prot <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(12){O}{D}DRx(6) OSODDRKENNDG	1	1:	Q9GU44 ck: 9332 len: 27 i Q9gu44 sycon raphanus. homeobox protein <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(20){T}DRx(4) RTSKYLAFTTRQALAAARLGLTDROVKV
1	1:	Q03977 ck: 4923 len: 19 i Q03977 escherichia coli. alpha-acceptor pol <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(13){T}DRx(3) MCMITDSLAVAVARTDRPSQ	1	1:	Q9XZ26 ck: 1592 len: 23 i Q9xz26 littorina saxatilis. calmodulin (c <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(8){A}DRx(11) LQDMINEVDADRDRLRPVPY
1	1:	Q99094 ck: 1642 len: 36 i Q99094 salmonella typhimurium. mkaa protein <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(15){D}DRx(18)	1	1:	

1	Q94681	ck: 5282	len: 25	1	Q94681	polyandrocampa misakiensis . pmhbox1
		<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N				
		x{20}{T}DRx{2}				
1:		HFNYLCRERERQEVAKAVNLDRQV				
1	Q23749	ck: 8475	len: 27	1	Q23749	ctenodrilus serratus . ovxl ortholog
		<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N				
		x{19}{T}DRx{5}				
1:		POEILVTDGTIAHRASPETDRSGEN				
1	Q25482	ck: 9269	len: 27	1	Q25482	metridium senile (brown sea anemone)
		<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N				
		x{20}{T}DRx{4}				
1:		HFNHFLTKERSEMATQNLMTDRQVKI				
1	Q94458	ck: 8578	len: 27	1	Q94458	chaetopterus variopedatus . chv-hb11
		<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N				
		x{20}{D}DRx{4}				
1:		NOKKFLKKDRDRISNEIGLDDRQIKY				
1	Q96892	ck: 9212	len: 27	1	Q96892	sacculina carcini . caudal (fragment)
		<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N				
		x{20}{T}DRx{4}				
1:		RFNMYITIKRKLELSRIILGUTDRQVKI				
1	P91901	ck: 9410	len: 27	1	P91901	homeobox protein bhox26
		<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N				
		x{20}{T}DRx{4}				
1:		LFNMVLTFRERLEISKSINLTDROVKI				
1	002602	ck: 9552	len: 27	1	002602	beroe ovata . homeobox protein bhox35
		<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N				
		x{20}{T}DRx{4}				
1:		LFNMVLTFRERLEISRGVNLTDROVKI				
1	018615	ck: 5276	len: 34	1	018615	artemia salina (brine shrimp) . ftz-f
		<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N				
		x{14}{A}DRx{17}				
1:		FKRTYONKKVYTCMADRSCHIDKSQRKRCPPGRF				
1	Q9V973	ck: 1645	len: 17	1	Q9V973	drosophila melanogaster (fruit fly) .
		<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N				
		x{4}{S}{T}DRx{9}				
1:		MELSTFDGPFHHAQOGR				
1	Q9TWRI	ck: 8106	len: 21	1	Q9TWRI	bombyx mori (silk moth) . cysteine pr
		<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N				
		x{5}{D}DRx{13}				
1:		XPEQVDDRKRGAVTDDKXXQX				
1	062575	ck: 7242	len: 26	1	062575	tetracapsula bryozoides . hox protein
		<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N				

1	Q957s8	ck: 1370	len: 23	i	Q957s8 pinus sylvestris (scots pine). ribos	1	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1	1:				x(17){T}Drx(3) MPTSNQSI RHGRKKRRTRDRTTRA	1:	x(11){T}Drx(15) VGCPRILMKCKTDRCCLTGCTCKRNGYCG
1	Q9T2K6	ck: 5922	len: 20	i	Q9T2K6 cucurbita moschata (cushaw squash)	1	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1	1:				x(3){S}{L}{A}Drx(12) GSSSLADRLXGSLAXDGFS	1:	x(11){D}Drx(2) NTNTQMTAYDQDDRGX
1	Q9T2H3	ck: 2853	len: 24	i	Q9T2H3 spinacia oleracea (spinach). chaperon	1	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1	1:				x(14){A}Drx(7) ATVVAPEKYSIKPTADRLIKIKE	1:	x(5){A}Drx(4) NLANKADRDQYE
1	Q37852	ck: 3123	len: 18	i	Q37852 bacteriophage r17. a protein (fragme	1	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1	1:				x(8){E}Drx(7) ALRYLALNEDRKFRSKHV	1:	x(6){T}Drx(8) MPVKKDDTRALSLEBY
1	Q42209	ck: 1580	len: 28	i	Q42209 arabidopsis thaliana (mouse-ear cress	1	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1	1:				x(8){T}Drx(17) PEIAHMYKTTDRAKYESTASWTQKYAMG	1:	x(4){T}Drx(20) MAENTDRNQTEKILNRYKLEQEVERL
1	Q9FZP4	ck: 3486	len: 29	i	Q9FZP4 arabidopsis thaliana (mouse-ear cress	1	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1	1:				x(20){E}Drx(6) MDVDVDFSTYLSKEIDENEDREPRKYV	1:	LFMNYLTRERLEISKITINLTDRQV
1	Q39633	ck: 1444	len: 23	i	Q39633 cucumis sativus (cucumber). catalase	1	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1	1:				x(4){A}Drx(16) NGSQADRSYGQKLAPHLVNRPST	1:	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1	P82195	ck: 7110	len: 26	i	P82195 spinacia oleracea (spinach). chloroph	1	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1	1:				x(6){E}Drx(17) KAHTRREDRTARHVRIRKKVECTPER	1:	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1	P82196	ck: 4154	len: 29	i	P82196 spinacia oleracea (spinach). chloroph	1	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1	1:				x(6){E}Drx(20) KAHTRREDRTARHVRIRKKVECTPERXXL	1:	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1	P83089	ck: 4307	len: 19	i	P83089 spinacia oleracea (spinach). thylakoid	1	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1	1:				x(12){T}Drx(4) AIAANAPLIDDTITTDRAVFFD	1:	MADEFIDYDRAVYSGFGGRC
1	Q9S8D2	ck: 2605	len: 29	i	Q9S8D2 cucumis melo (muskmelon). cmet1-b=tr	1	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1	1:					1:	MSEPKAIDPKLSTTDRAVVK

```

1      062256  ck: 3193  len: 18      i 062256 mus musculus (mouse). spermatogenic-
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(10){E}DRx(5)
1      MSSGKODSFWEDRIPPGR
1      063985  ck: 5721  len: 25      i 063985 rattus sp. hox1.8 homeobox homolog F
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(20){T}DRx(2)
1      LFNMVLTFRERRLRLEISRSVHLTDROY
1      061159  ck: 837   len: 32      i 061159 mus musculus (mouse). m8-ag (fragme
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(15){A}DRx(14)
1      MAASMCDFSFSCVGVADRARGSVVEVRYDSIK
1      0923H1  ck: 2032  len: 40      i 0923h1 cricetus griseus (chinese hamster)
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(19){A}DRx(16)
1      PTPAVEGAVASGALADPATTAADRASSIAAXGSRPRSM
1      099KX5  ck: 9106  len: 27      i 099KX5 mus musculus (mouse). hypothetical 2
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(12){A}DRx(12)
1      MASLPGVGSQVPADRGYIGPPRAVOEH
1      09QV00  ck: 3385  len: 24      i 09QV00 cavia (guinea pigs). vitronectin (fr
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(14){A}DRx(7)
1      XXESXKGRXTBGFNADRKXQXXEL
1      09QV89  ck: 581   len: 32      i 09QV89 rattus sp. s1p-14=fatty acid-binding
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(14){D}DRx(14)
1      EISPOLGVEFDEVTAADRKXSVVTLDGKLY
1      09QV71  ck: 1363  len: 17      i 09QV71 rattus sp. lactase-phlorizin hydrola
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      (D){W}{E}DRx(12)
1      DMEDNRFTIAGFLTNDL
1      061461  ck: 4925  len: 25      i 061461 mus musculus (mouse). cytochrome p-4
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(17){D}DRx(15)
1      SHRLPTLDDRSKMPYTDVHIEIOR
1      065747  ck: 4484  len: 19      i 065747 bluetongue virus. outer coat protein
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(13){E}DRx(2)
1      ERLKTFEHRNQRRDEDRRY
1      066538  ck: 8137  len: 35      i 066538 ebola virus. 3' proximal protein (fd
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(12){D}DRx(20)
1      MRKINNFSLKFDNRMLKLIKILICNHTVDESPHYS
1      083622  ck: 3517  len: 9       i 083622 murray valley encephalitis virus.
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(4){T}DRx(2)
1      THVSEDRV
1      084254  ck: 9066  len: 22      i 084254 bovine papillomavirus. x protein
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(13){A}DRx(6)
1      NAGPKGTTPEDVADRPDLPE
1      09PWC1  ck: 9645  len: 39      i 09PWC1 brachydanio rerio (zebrafish) (ze
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(19){D}DRx(17)
1      LVELSVLPSQLQPHVESPDNRSLHAMFGERSLSLYK
1      093438  ck: 3493  len: 29      i 093438 gallus gallus (chicken). lhx1 pro
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(10){A}DRx(16)
1      RFLXOKTLPADRDQIAQOLGILTNAQVIT
1      090297  ck: 9410  len: 27      i 090297 carassius auratus (goldfish). hom
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(20){T}DRx(4)
1      LFMNVLFRERRLRLEISRSINLTDROYKI
1      090298  ck: 9642  len: 27      i 090298 carassius auratus (goldfish). hom
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(20){T}DRx(4)
1      LFMNVLFRERRLRLEISRSINLTDROYKI
1      007145  ck: 9552  len: 27      i 007145 petromyzon marinus (sea lamprey).
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(20){T}DRx(4)
1      LFMNVLFRERRLRLEISRSINLTDROYKI
1      007150  ck: 9407  len: 27      i 007150 petromyzon marinus (sea lamprey).
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(20){T}DRx(4)
1      LFMNVLFRERRLRLEISRSINLTDROYKI
1      073591  ck: 8165  len: 14      i 073591 gallus gallus (chicken). hox c10
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(10){T}DRx
1      RLKISKSINLTDRO
1      P83009  ck: 7125  len: 13      i P83009 lamna nasus. phospholipase (fxu
1      <X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
1      x(6){N}{D}DRx(3)

```

```

Q9KLI3 ck: 1300 len: 32 | Q9KLI3 vibrio ch.
1:
<x{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,,
x{13})(T)DRx{16}
MPDRLRFSATCSTDRRKQTNLFSPLKQOMR

08X407 ck: 6940 len: 13 | 08X407 escherichia coli o157:h7. hypothe
1:
<x{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,
x{5})(T)DRx{8}
MSTRKEFVMLLFH

Q8VJ09 ck: 8356 len: 35 | Q8VJ09 mycobacterium tuberculosis. hypot
1:
<x{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,
x{13})(A)DRx{19}
MKYKLAIDEDYDRADRTERGAILRENNLYSSILTE

Q9HNL3 ck: 9872 len: 32 | Q9HNL3 halobacterium sp. (strain nrc-1).
1:
<x{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,
x{17})(A)DRx{12}
MVRVPTVDAGRAEMRERADRLETTAFAWRLYVD

Databases searched:
SWISS-PROT, Release 40.3, Released on 9Aug2002, Formatted on 20Aug2002
SPTREMBL, Release 21.0, Released on 15Jun2002, Formatted on 28Jun2002

Total finds: 118
Total length: 247,523,443
Total sequences: 784,472
CPU time: 04:50.37

```

```
Databases searched:
  SWISS-PROT, Release 40.3, Released on 9Aug2002, Formatted on 20Aug2002
  SPTSRMBL, Release 21.0, Released on 15Jun2002, Formatted on 28Jun2002

Total finds:      118
Total length:    247,523,443
Total sequences: 784,472
CPU time:        04:50.37
```



```

1 FINDPATTERNS on pir:* allowing 0 mismatches
1 1 <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N,M,P,V,W){0,1}
A21630 ck: 4925 len: 25 1 cytochrome P450b - mouse (fragment)
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N
x{7}(D)DRX{15}
SHRLPTLDDRSKMPYTDVAVHETIQ
F42762 ck: 5572 len: 20 1 C.3.4.25.1 proteasome endopeptidase complex
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N
x{9}(A)DRX{7}
MKGMXVATADRRFGIOAQ
S55729 ck: 1645 len: 28 1 orotidine-5'-monophosphate decarboxylase -
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N
x{16}(A)DRX{9}
VTVSADVTTKELDLADRVMGFVSTRS
B44336 ck: 2946 len: 34 1 neurotoxin Tx3-2 - spider (Phonetreria nigri
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N
x{17}(E)DRX{14}
ACAGLYKKCKGKSGSPCEDRPCDLMGNCIC
B28563 ck: 9362 len: 22 1 hemoglobin chain IV - earthworm (Lumbricus
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N
x{9}(E)DRX{10}
ADESCSCSYEDRRVYHIMDDV
S01808 ck: 3406 len: 24 1 hemoglobin AII - tube worm (Lamellibrachia
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N
x{7}(E)DRX{14}
SSNSCTEDRREMQLMANYMSAQ
D60894 ck: 6311 len: 20 1 gamma crystallin V - bullfrog (fragment)
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N
x{6}(E)DRX{11}
GKIVFIEDRNFQGRSECCS
S04621 ck: 5853 len: 20 1 gamma-crystallin II - milk shark (fragment)
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N
x{6}(E)DRX{11}
GKITFYEDRGFOGHCYECSS
A61570 ck: 6107 len: 20 1 gamma-crystallin (total) - siamang (fragment)
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N
x{6}(E)DRX{11}
GKITFYEDRGFOGRHYECTT
B61570 ck: 6068 len: 20 1 gamma-crystallin (total) - Formosa sika deer
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N
x{6}(E)DRX{11}
GKITFYEDRGFOGRHYECSS
A60502 ck: 4514 len: 25 1 myonexin - northern leopard frog (fragme
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,
x{11}(E)DRX{11}
EDGDFEEDYDGEDRVIFISLKNYKA
E33208 ck: 2821 len: 29 1 calreticulin, uterine - rabbit (fragment)
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,
x{16}(T)DRX{10}
EPVYVFKFQFLDGDGWTDRWIESKHKSDF
C33208 ck: 2871 len: 29 1 calreticulin, slow twitch skeletal muscle
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,
x{16}(T)DRX{10}
EPVYVFKFQFLDGDGWTDRWIESKHKSDF
A33434 ck: 2817 len: 24 1 calcium-binding protein - dog (fragment)
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,
x{16}(T)DRX{5}
EPATYFKFQFLDGDGXTDRXIESK
B41481 ck: 1642 len: 36 1 virulence-associated protein (vira 5' re
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,
x{15}(D)DRX{18}
GSQDFASQSLKRLSDPDRTADTRIKRIIMRVLNS
S07394 ck: 9004 len: 22 1 DNA-invertase - phage D108 (fragment)
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,
x{13}(D)DRX{6}
YKKHPARKTHIENDRINOIDR
S35552 ck: 9008 len: 32 1 vesicle-associated membrane protein 1 -
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,
x{9}(D)DRX{20}
ERDQKLSLDDRADALQAGASQFESSAKILKR
S35553 ck: 9118 len: 32 1 vesicle-associated membrane protein 1 -
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,
x{9}(D)DRX{20}
ERDQKLSLDDRADALQAGASQFESSAKILKR
S35555 ck: 9033 len: 32 1 vesicle-associated membrane protein 2 -
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,
x{9}(D)DRX{20}
ERDQKLSLDDRADALQAGASQFESSAKILKR
S35554 ck: 9033 len: 32 1 vesicle-associated membrane protein 2 -
1: <X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,
x{9}(D)DRX{20}
ERDQKLSLDDRADALQAGASQFESSAKILKR
E44621 ck: 9552 len: 27 1 homeotic protein Hox 10 (clone 10w) - se

```

1			<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(20){T}Drx(4) LFNMVLTFRRLRLEISGCVNLTDROVKI	1:	MTTDRAGLGROLOMIRGLH		
1	F44621	ck: 9407 len: 27	i homeotic protein Hox 10 (clone 10x) - sea 1	1	S66213	ck: 1672 len: 17	i glucose 1-dehydrogenase (EC 1.1.1.47) -
			<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(20){T}Drx(4) LFSMVLTRRLRLRLEISHLSLTDROVKI	1:	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(9){E}Drx(5) MKALAVKRGEDRPVIVIE		
1	C84355	ck: 9872 len: 32	i hypothetical protein Vng2049c [imported] -	1	S23180	ck: 7956 len: 31	i carboxypeptidase - Sulfolobus solfataric
			<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(17){A}Drx(12) MVRVPVTDAGRAEMRERADRLTETAAFWRLVD	1:	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(14){E}Drx(14) MDLVEKLKNDVKEIEDRIIQIRKIXENPPL		
1	E64634	ck: 575 len: 23	i hypothetical protein HP0917 - Helicobacter	1	S29326	ck: 9119 len: 22	i hypothetical protein 22, psba 5'-region
			<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(13){E}Drx(7) MSPLTFLRNPFLTQEDRFEQEIITA	1:	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(5){T}Drx(14) MEVLTDRSIECGIYLRKIESI		
1	PH0858	ck: 807 len: 23	i Maud protein - Paracoccus denitrificans (fr	1	T10123	ck: 1444 len: 23	i probable catalase (EC 1.11.1.6) - cucumb
			<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N (L){E}{A}Drx(18) LEADRTGFASLQGYMASRRKQAA	1:	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(4){A}Drx(16) NGSQADRSVCGKLAFLHNVPSI		
1	A61412	ck: 8072 len: 27	i methane monooxygenase (EC 1.14.13.25) prote	1	PS0212	ck: 8963 len: 15	i 29k protein 4228 - rice (strain Nihonbar
			<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(12){A}Drx(12) WIEDYASRIDFKADRDQIVKAVLAGLK	1:	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(2){D}Drx(9) SPADDRRDVGDRIYAD		
1	A82288	ck: 3787 len: 33	i hypothetical protein VC0735 [imported] - Vi	1	B60698	ck: 2526 len: 29	i trichocyst protein 27 - Paramecium tetra
			<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(15){A}Drx(15) MPVGSSIFPCISIFALADKRVSVRYWCFFTLVL	1:	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(13){E}Drx(13) DPLDRLLSTLTDLDRVYVAEQKEDDAKNO		
1	B82421	ck: 1300 len: 32	i hypothetical protein VCA0761 [imported] - V	1	S19614	ck: 1820 len: 17	i globin - polychaete (Eudistylia vancouver
			<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(13){T}Drx(16) MPDRLLRESAICSTDRRKQTNLFSSDLKQOWR	1:	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(7){E}Drx(7) GMKXXSMEDRKTVLADW		
1	B39089	ck: 6231 len: 20	i hydrogenase (EC 1.18.99.1) 34k chain - Thic	1	A60741	ck: 4352 len: 25	i insulin-like growth factor-binding prote
			<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(4){A}Drx(13) XXEQADRPVSIGLPPQEXTG	1:	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(20){E}Drx(2) LAPGPGGVQAGKPGXVEEDRGC		
1	H70954	ck: 8577 len: 27	i hypothetical protein Rv3599c - Mycobacteriu	1	B28457	ck: 4360 len: 19	i proteoglycan II, bone - human (tentative
			<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N x(11){A}Drx(12) MPASSLGTGSPADRDLATHERRREVI	1:	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(11){D}Drx(5) DEAXGIAPEVPDDRPEPS		
1	S14161	ck: 5957 len: 20	i probable dTDP-4-keto-6-deoxy-hexose 3,4-isc	1	C54037	ck: 1565 len: 29	i splicing regulatory protein SWAP homolog
			x(3){T}Drx(14)	1:	<X(0,20){A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F, x(17){E}Drx(9) POLTOBELEAKQAKKLEDRLLAAAREKL		

1	A48845	ck: 9975	len: 23	sterol regulatory element 1 binding protein	1:	x(S)(S)(A)DRx ASSDADRG
1					1:	
1					1:	
1	S22228	ck: 4810	len: 25	vitronectin - dog (fragment)	1:	x(S)(S)(A)DRx ASGEDRG
1					1:	
1	A26393	ck: 7	len: 16	annexin 36k chain - pig (fragment)	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	S68260	ck: 6375	len: 38	hypothetical protein gadd7.1 - long-tailed	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	A35678	ck: 3193	len: 18	hypothetical protein (proenkephalin 5' regi	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	PH0780	ck: 8887	len: 15	T-cell receptor alpha chain (C11) - mouse	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	PH0794	ck: 1531	len: 17	T-cell receptor alpha chain (K1 V-alpha-3.4	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	PH0806	ck: 8864	len: 15	T-cell receptor alpha chain (P5.1.1V-alpha	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	S03505	ck: 6013	len: 20	T-cell receptor alpha chain J region (80)	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	PT0212	ck: 3993	len: 10	T-cell receptor alpha chain V-J region (4-1	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	PT0547	ck: 2627	len: 8	T-cell receptor beta chain V-D-J region (12	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	PT0676	ck: 2049	len: 7	T-cell receptor beta chain V-D-J region	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	PT0576	ck: 2172	len: 7	T-cell receptor beta chain V-D-J region	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	PT0366	ck: 480	len: 28	T-cell receptor beta chain V-J region (6	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	A46592	ck: 1363	len: 17	lactase-phlorizin hydrolase, 200k isoform	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	S78414	ck: 6086	len: 26	ribosomal protein RL25, mitochondrial [v	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	S22227	ck: 5585	len: 25	vitronectin - guinea pig (fragment)	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	A54226	ck: 6170	len: 26	light-harvesting protein B-830 alpha-1 c	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	B54226	ck: 634	len: 23	light-harvesting protein B-830 alpha-2 c	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	T46593	ck: 6757	len: 38	phytoene dehydrogenase [imported] - Myco	1:	x(S)(S)(A)DRx ASSDORT
1					1:	
1	H85575	ck: 6940	len: 13	hypothetical protein Z0899 [imported] -	1:	x(S)(S)(A)DRx ASSDORT
1					1:	

1 PA0041 ck: 9117 len: 15 ! plastoquinol-plastocyanin reductase (EC 1.1.1.1)
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N
x{5})(A)DRX(7)
1: ASSIPADRVPDMEKR

Databases searched:
NBRF, Release 73.0, Released on 16Aug2002, Formatted on 20Aug2002

Total finds: 63
Total length: 96,134,422
Total sequences: 283,224
CPU time: 01:39.36

```

11AA_SEQUENCE 1.0
F1:A21630 - cytochrome P450b - mouse (fragment)
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Mus musculus (house mouse)
C:Accession: B44336
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 11-Jun-1999
R:Stupans, I.; Ikeda, T.; Kessler, D.J.; Nebert, D.W.
DNA 3, 129-137, 1984
A:Title: Characterization of a cDNA clone for mouse phenobarbital-inducible
cytochrome P-450b.
A:Reference number: A21630; MUID:84207435; PMID:6547088
A:Accession: A21630
A:Molecule type: mRNA
A:Residues: 1-25 <STUD>
A:Cross-references: BG:K02409; NID:9192896; PIDN:AAA37510.1; PID:9553905
A>Note: The authors translated the codon CTA for residue 5 as Pro, ACC for
residue 7 as Ser and AGT for residue 12 as Thr
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: electron transfer; heme; monooxygenase; oxidoreductase;
transmembrane protein

A21630 Length: 25 December 22, 2002 19:21 Type: P Check: 4925 ..

1 SHRPLTLDDR SKMPYTDAYI HEIQR

11AA_SEQUENCE 1.0
F1:F42762 - C 3.4.25.1 proteasome endopeptidase complex ( ) subunit 13 - bovine
(fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 03-Jun-2002
C:Accession: F42762
R:Diick, L.R.; Mooney, C.R.; Pramanik, B.C.; DeMartino, G.N.; Slaughter, C.A.
Biochemistry 31, 7347-7355, 1992
A:Title: Identification and localization of a cysteinyl residue critical for
the trypsin-like catalytic activity of the proteasome.
A:Reference number: A42762; MUID:92378961; PMID:1510924
A:Accession: F42762
A>Status: Preliminary
A:Molecule type: protein
A:Residues: 1-20 <DIC>
A>Note: sequence extracted from NCBI backbone (NCBI:P112175)
C:Superfamily: multicatalytic endopeptidase complex chain C9
C:Keywords: hydrolase

F42762 Length: 20 December 22, 2002 19:21 Type: P Check: 5572 ..

1 MKGNXVAIA ADRFGIQAO

11AA_SEQUENCE 1.0
F1:S55729 - orotidine-5'-monophosphate decarboxylase - Aspergillus awamori
(fragment)
C:Species: Aspergillus awamori
C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C:Accession: S55729
R:Gouka, R.J.; Hessing, J.G.M.; Stam, H.; Musters, W.; van den Hondel,
C.A.M.J.J.
Curr. Genet. 27, 536-540, 1995
A:Title: A novel strategy for the isolation of defined pyrG mutants and the
development of a site-specific integration system for Aspergillus awamori.
A:Reference number: S55729; MUID:96031709; PMID:7553938
A:Accession: S55729
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-18,19-28 <GOU>
C:Superfamily: orotidine-5'-phosphate decarboxylase; orotidine-5'-phosphate
decarboxylase homology

S55729 Length: 28 December 22, 2002 19:21 Type: P Check: 1645 ..

1 VTVSADVTTT KELDLADRV MGFVSTRS

11AA_SEQUENCE 1.0
F1:B44336 - neurotoxin Tx3-2 - spider (Phoneutria nigriventer)
C:Species: Phoneutria nigriventer
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Aug-1998
C:Accession: B44336
R:Cordeiro M do, N.; de Figueiredo, S.G.; Valentim A do, C.; Diniz, C.R.; von
Eickstedt, V.R.; Gilroy, J.; Richardson, M.
Toxicol 31, 35-42, 1993
A:Title: Purification and amino acid sequences of six Tx3 type neurotoxins from
the venom of the Brazilian 'armed' spider Phoneutria nigriventer (Keys.).
A:Reference number: A44336; MUID:93190315; PMID:8446961
A:Accession: B44336
A>Status: Preliminary
A:Molecule type: protein
A:Residues: 1-34 <COR>
A:Experimental source: venom
A>Note: sequence extracted from NCBI backbone (NCBI:P126773)
C:Superfamily: cutatoxin
C:Keywords: neurotoxin

B44336 Length: 34 December 22, 2002 19:21 Type: P Check: 2946 ..

1 ACAGLYKKKG KGASPCEDR PCKDLAMGN CICK

11AA_SEQUENCE 1.0
F1:B28563 - hemoglobin chain IV - earthworm (Lumbricus terrestris) (fragment)
C:Species: Lumbricus terrestris (common earthworm)
C>Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 04-Mar-2000
C:Accession: B28563
R:Goroch, T.; Shishikura, F.; Snow, J.W.; Erefef, K.I.; Vinogradov, S.N.; Walz,
D.A.
Biochem. J. 241, 441-445, 1987
A:Title: Two globin strains in the giant annelid extracellular haemoglobins.
A:Reference number: A90337; MUID:87241210; PMID:3593201
A:Accession: B28563
A:Molecule type: protein
A:Residues: 1-22 <GOT>
C:Superfamily: globin; globin homology
C:Keywords: oxygen carrier

B28563 Length: 22 December 22, 2002 19:21 Type: P Check: 9362 ..

1 ADEBSCSTE DREVRHWD DV

11AA_SEQUENCE 1.0
F1:S01808 - hemoglobin AII - tube worm (Lamellibrachia sp.) (fragment)
C:Species: Lamellibrachia sp.
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 04-Mar-2000
C:Accession: S01808
R:Suzuki, T.; Takagi, T.; Ohta, S.
Biochem. J. 255, 541-545, 1988
A:Title: N-terminal amino acid sequence of the deep-sea tube worm haemoglobin.
remarkably resembles that of annelid haemoglobin.
A:Reference number: S01807; MUID:89076216; PMID:3202832
A:Accession: S01808
A:Molecule type: protein
A:Residues: 1-24 <SUZ>
C:Superfamily: globin; globin homology
C:Keywords: oxygen carrier

S01808 Length: 24 December 22, 2002 19:21 Type: P Check: 3406 ..

1 SSNSCTEDR REMOLMANV WSAO

11AA_SEQUENCE 1.0
F1:D60894 - gamma crystallin V - bullfrog (fragment)
C:Species: Rana catesbeiana (bullfrog)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Nov-1997
C:Accession: D60894
R:Chiou, S.H.
J. Protein Chem. 7, 527-534, 1988
A:Title: The protein sequence homology of gamma-crystallins among major
vertebrate classes and their DNA sequence homology to heat-shock protein genes.
A:Reference number: A60894; MUID:89351593; PMID:3253376
A:Accession: D60894

```

A:Molecule type: protein
A:Residues: 1-20 <CHI>
C:Superfamily: beta-crystallin
C:Keywords: duplication; eye lens

D60894 Length: 20 December 22, 2002 19:21 Type: P Check: 6311 ..

1 GKTFYEDRG FQGRSYECSS

!!AA_SEQUENCE 1.0

F1:S04621 - gamma-crystallin II - milk shark (fragment)

C:Species: Rhizoprionodon acutus (milk shark)

C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Nov-1997

C:Accession: S04621

R:Chiou, S.H.

FERS lett. 250, 25-29, 1989

A:Title: Characterization of gamma-crystallins from eye lenses of shark: closer structural similarity to mammalian than other piscine gamma-crystallins?

A:Reference number: S04621; MUID:89290026; PMID:2737298

A:Accession: S04621

A:Molecule type: protein

A:Residues: 1-20 <CHI>

C:Superfamily: beta-crystallin

C:Keywords: duplication

S04621 Length: 20 December 22, 2002 19:21 Type: P Check: 5853 ..

1 GKTFYEDRG FQGHYECSS

!!AA_SEQUENCE 1.0

F1:A61570 - gamma-crystallin (total) - siamang (fragment)

C:Species: Hylobates syndactylus (siamang)

C>Date: 25-Oct-1994 #sequence_revision 06-Jan-1995 #text_change 21-Nov-1997

C:Accession: A61570

R:Chiou, S.H.; Chang, W.P.; Ting, L.M.; Lai, T.A.; Lin, H.K.

Curr. Eye Res. 7, 1017-1022, 1988

A:Title: Biochemical characterization of lens crystallins from three mammalian species.

A:Reference number: A61570; MUID:89152306; PMID:3229121

A:Accession: A61570

A:Molecule type: protein

A:Residues: 1-20 <CHI>

A>Note: 8-Glu, 9-Gln, 10-Ala, 15-Ser, 18-Thr, and 20-Ser were also found

A>Note: this preparation appears to represent the product of several gamma crystallin genes from one specimen

C:Superfamily: beta-crystallin

C:Keywords: duplication; eye lens

A61570 Length: 20 December 22, 2002 19:21 Type: P Check: 6107 ..

1 GKTFYEDRG FQGRHYECCT

!!AA_SEQUENCE 1.0

F1:B61570 - gamma-crystallin (total) - Formosa sika deer (fragment)

C:Species: Cervus nippon taiouanus (Formosa sika deer)

C>Date: 25-Oct-1994 #sequence_revision 06-Jan-1995 #text_change 21-Nov-1997

C:Accession: B61570

R:Chiou, S.H.; Chang, W.P.; Ting, L.M.; Lai, T.A.; Lin, H.K.

Curr. Eye Res. 7, 1017-1022, 1988

A:Title: Biochemical characterization of lens crystallins from three mammalian species.

A:Reference number: A61570; MUID:89152306; PMID:3229121

A:Accession: B61570

A:Molecule type: protein

A:Residues: 1-20 <CHI>

A>Note: 6-Phe, 14-His, and 15-Cys were also found

A>Note: this preparation appears to represent the product of several gamma crystallin genes from one specimen

C:Superfamily: beta-crystallin

C:Keywords: duplication; eye lens

B61570 Length: 20 December 22, 2002 19:21 Type: P Check: 6068 ..

1 GKTFYEDRG FQGRHYECSS

!!AA_SEQUENCE 1.0

F1:A60502 - myonexin - northern leopard frog (fragment)

C:Species: Rana pipiens (northern leopard frog)

C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999

C:Accession: A60502; A33087

R:Urball, J.G.

Dev. Biol. 142, 103-114, 1990

A:Title: Myonexin: an 80-kDa glycoprotein that binds fibronectin and is located at embryonic myotendinous junctions.

A:Reference number: A60502; MUID:91032554; PMID:2146177

A:Accession: A60502

A:Molecule type: protein

A:Residues: 1-25 <TID>

C:Comment: This protein resembles calsequestrin in sequence but differs in distribution and physical properties.

C:Superfamily: calsequestrin

C:Keywords: fibronectin binding; glycoprotein; skeletal muscle

A60502 Length: 25 December 22, 2002 19:21 Type: P Check: 4514 ..

1 EDGFDFPEYD GEDRVIFISL KNYKA

!!AA_SEQUENCE 1.0

F1:E33208 - calreticulin, uterine - rabbit (fragment)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 31-Jul-1991 #sequence_revision 31-Jul-1991 #text_change 01-Mar-1996

C:Accession: E33208; F33208

R:Miller, R.E.; Baksh, S.; Shemanko, C.; Carpenter, M.R.; Smillie, L.; Vance, J.E.; Opsas, M.; Michalak, M.

J. Biol. Chem. 266, 7155-7165, 1991

A:Title: Calreticulin, and not calsequestrin, is the major calcium binding

protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic reticulum.

A:Reference number: A33208; MUID:91201375; PMID:2016321

A:Accession: E33208

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-29 <MIT>

A>Note: uterine form

A:Accession: F33208

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-20 <MIT>

A>Note: hepatic form

C:Superfamily: calreticulin

E33208 Length: 29 December 22, 2002 19:21 Type: P Check: 2821 ..

1 EPVYFQFQF LDGDGWTDRW IESKHSDF

!!AA_SEQUENCE 1.0

F1:C33208 - calreticulin, slow twitch skeletal muscle - rabbit (fragment)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 31-Jul-1991 #sequence_revision 31-Jul-1991 #text_change 07-Feb-1997

C:Accession: C33208

R:Miller, R.E.; Baksh, S.; Shemanko, C.; Carpenter, M.R.; Smillie, L.; Vance, J.E.; Opsas, M.; Michalak, M.

J. Biol. Chem. 266, 7155-7165, 1991

A:Title: Calreticulin, and not calsequestrin, is the major calcium binding

protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic reticulum.

A:Reference number: A33208; MUID:91201375; PMID:2016321

A:Accession: C33208

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-29 <MIT>

C:Superfamily: calreticulin

C:Keywords: skeletal muscle

C33208 Length: 29 December 22, 2002 19:21 Type: P Check: 2871 ..

1 EPVYFQFQF LDGDGWTDRW IESKHSDF

```
!!AA_SEQUENCE 1.0
FI:A3334 - calcium-binding protein - dog (fragment)
C:Species: Canis lupus familiaris (dog)
C>Date: 21-Feb-1990 #sequence_revision 21-Feb-1990 #text_change 16-Feb-1997
C:Accession: A3334
R:Collins, J.H.; XL, Z.; Alderson-Lang, B.H.; Treves, S.; Volpe, P.
Biochem. Biophys. Res. Commun. 164, 575-579, 1989
A>Title: Sequence homology of a canine brain calcium-binding protein with
calregulin and the human Ro/SS-A antigen.
A:Reference number: A3334; MUID:90026437; PMID:2803321
A:Accession: A3334
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-24 <COL>
C:Superfamily: calreticulin
C:Keywords: calcium binding

A33434 Length: 24 December 22, 2002 19:21 Type: P Check: 2817 ..

1 EPAIYFEQF LDGDXTRX IESK

!!AA_SEQUENCE 1.0
FI:B4181 - virulence-associated protein (vira 5' region) - Salmonella
typhimurium (fragment)
C:Species: Salmonella typhimurium
C>Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 20-Aug-1999
C:Accession: B4181
R:Gullig, P.A.; Chiodo, V.A.
Infect. Immun. 58, 2651-2658, 1990
A>Title: Genetic and DNA sequence analysis of the Salmonella typhimurium
virulence plasmid gene encoding the 28,000-molecular-weight protein.
A:Reference number: A41481; MUID:90316693; PMID:2164511
A:Accession: B41481
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-36 <COL>
A:Cross-references: GB:M64295; GB:M33927; NID:g154440; PID:AAA27272.1;
PID:g154441
C:Superfamily: virulence-associated protein spvB
C:Keywords: DNA binding; transcription regulation

B41481 Length: 36 December 22, 2002 19:21 Type: P Check: 1642 ..

1 GSOFFASQLS KLRISDDRTA DTRIKRIIN MRVLNS

!!AA_SEQUENCE 1.0
FI:S07394 - DNA-Invertase - phage D108 (fragment)
C:Species: phage D108
A:Note: host Escherichia coli
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Jun-2000
C:Accession: S07394
R:Szatmari, G.B.; Lapointe, M.; Dubow, M.S.
Nucleic Acids Res. 15, 6691-6704, 1987
A>Title: The right end of transposable bacteriophage D108 contains a 520 base
pair protein-encoding sequence not present in bacteriophage M1.
A:Reference number: S07394; MUID:87316928; PMID:2857646
A:Accession: S07394
A:Molecule type: DNA
A:Residues: 1-22 <SZA>
A:Cross-references: EMBL:X05926; NID:g14918; PID:CAA29365.1; PID:g1364174
C:Genetics:
A:Gene: gin
C:Superfamily: transposase repressor
C:Keywords: DNA binding; DNA integration; DNA recombination

S07394 Length: 22 December 22, 2002 19:21 Type: P Check: 9004 ..

1 YKKHPAKRTH IENDRINOI DR

!!AA_SEQUENCE 1.0
FI:S3552 - vesicle-associated membrane protein 1 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 10-Dec-1993 #sequence_revision 27-Feb-1997 #text_change 17-Mar-1999
```

```
C:Accession: S3552
R:Patarnello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.
Nature 364, 581-582, 1993
A>Title: Neurotransmission and secretion.
A:Reference number: S3552; MUID:93354436; PMID:8350916
A:Accession: S3552
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-32 <PAT>
C:Superfamily: synaptobrevin

S3552 Length: 32 December 22, 2002 19:21 Type: P Check: 9008 ..

1 ERDQKLESD DRADALQAGA SFESSAKL KR

!!AA_SEQUENCE 1.0
FI:S3553 - vesicle-associated membrane protein 1 - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 20-May-1994 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C:Accession: S3553
R:Patarnello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.
Nature 364, 581-582, 1993
A>Title: Neurotransmission and secretion.
A:Reference number: S3552; MUID:93354436; PMID:8350916
A:Accession: S3553
A:Molecule type: DNA
A:Residues: 1-32 <PAT>
C:Superfamily: synaptobrevin

S3553 Length: 32 December 22, 2002 19:21 Type: P Check: 9118 ..

1 ERDQKLESD DRADALQAGA SFESSAKL KR

!!AA_SEQUENCE 1.0
FI:S3555 - vesicle-associated membrane protein 2 - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 10-Dec-1993 #sequence_revision 26-Jul-1996 #text_change 17-Mar-1999
C:Accession: S3555
R:Patarnello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.
Nature 364, 581-582, 1993
A>Title: Neurotransmission and secretion.
A:Reference number: S3552; MUID:93354436; PMID:8350916
A:Accession: S3555
A:Molecule type: DNA
A:Residues: 1-32 <PAT>
C:Superfamily: synaptobrevin
C:Keywords: membrane trafficking; transmembrane protein

S3555 Length: 32 December 22, 2002 19:21 Type: P Check: 9033 ..

1 ERDQKLESD DRADALQAGA SFESSAKL KR

!!AA_SEQUENCE 1.0
FI:S3554 - vesicle-associated membrane protein 2 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 10-Dec-1993 #sequence_revision 27-Feb-1997 #text_change 17-Mar-1999
C:Accession: S3554
R:Patarnello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.
Nature 364, 581-582, 1993
A>Title: Neurotransmission and secretion.
A:Reference number: S3552; MUID:93354436; PMID:8350916
A:Accession: S3554
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-32 <PAT>
C:Superfamily: synaptobrevin

S3554 Length: 32 December 22, 2002 19:21 Type: P Check: 9033 ..

1 ERDQKLESD DRADALQAGA SFESSAKL KR

!!AA_SEQUENCE 1.0
FI:E44621 - homeotic protein Hox 10 (clone 10w) - sea lamprey (fragment)
```

```

C:Species: Petromyzon marinus (sea lamprey)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 15-Oct-1999
C:Accession: E44621
R:Pendleton, J.W.; Nagai, B.K.; Murtha, M.T.; Ruddle, F.H.
Proc. Natl. Acad. Sci. U.S.A. 90, 6300-6304, 1993
A:Title: Expansion of the Hox gene family and the evolution of chordates.
A:Reference number: A44616; MUID:93317669; PMID:8101001
A:Accession: E44621
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-27 <PEN>
A:Cross-references: GB:I14895; NID:9290813; PIDN:AAA02530.1; PID:9290814
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation

E44621 Length: 27 December 22, 2002 19:21 Type: P Check: 9552 ..

1 LFMNLTFRER RLEISRGVNL TDRQVKI

!IAA_SEQUENCE 1.0
F1:E44621 - homeotic protein Hox 10 (clone 10x) - sea lamprey (fragment)
C:Species: Petromyzon marinus (sea lamprey)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 15-Oct-1999
C:Accession: E44621
R:Pendleton, J.W.; Nagai, B.K.; Murtha, M.T.; Ruddle, F.H.
Proc. Natl. Acad. Sci. U.S.A. 90, 6300-6304, 1993
A:Title: Expansion of the Hox gene family and the evolution of chordates.
A:Reference number: A44616; MUID:93317669; PMID:8101001
A:Accession: E44621
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-27 <PEN>
A:Cross-references: GB:I14900; NID:9290819; PIDN:AAA02535.1; PID:9290820
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation

E44621 Length: 27 December 22, 2002 19:21 Type: P Check: 9407 ..

1 LFSMYLTRER RLEISHLTSL TDRQVKI

!IAA_SEQUENCE 1.0
P1:G84355 - hypothetical protein Vng2049c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84355
R:Ng, W.V.; Kennedy, S.P.; Mahaitas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.;
Lasky, S.R.; Balfag, N.; Thorsson, V.; Strogna, J.; Swartzell, S.; Weit, D.;
Hall, J.; Dahl, T.A.; Weili, R.; Goo, Y.A.; Lelhausert, B.; Keller, K.; Cruz,
R.; Danson, M.D.; Hough, D.W.; Maddocks, D.G.; Jablonski, P.E.; Krebs, M.P.;
Angelyne, C.M.; Dale, H.; Isenbarger, T.A.; Peck, R.F.; Pohlschrod, M.;
Spudis, J.L.; Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe,
T.M.; Liang, P.; Riley, M.; Hood, L.; Dassarma, S.
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: C84355
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-32 <STO>
A:Cross-references: GB:AE004437; NID:g10581474; PIDN:AMG20207.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNG2049C

C84355 Length: 32 December 22, 2002 19:21 Type: P Check: 9872 ..

1 MVRVPVTDAG RAEMREERADR LETTAAPWRL VD

!IAA_SEQUENCE 1.0
P1:E64634 - hypothetical protein HP0917 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C:Accession: E64634

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.;
Fleischmann, R.D.; Ketchum, K.A.; Klein, H.P.; Gill, S.; Dougherty, B.A.;
Nelson, K.; Quackenbush, J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus,
B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, K.;
Fitzgerald, L.M.; Lee, N.; Adams, M.D.; Adams, E.K.; Berg, D.E.; Gocayne,
J.D.; Ullrich, T.R.; Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Meldman,
J.M.; Fujii, C.; Bowman, C.; Wathey, L.
Nature 388, 539-547, 1997
A:Authors: Mallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.;
Fraser, C.M.; Venter, J.C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter
pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: E64634
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-23 <TOM>
A:Cross-references: GB:AE000601; GB:AE000511; NID:g2314051; PIDN:AAD07968.1;
PID:g2314059; TIGR:HP0917

E64634 Length: 23 December 22, 2002 19:21 Type: P Check: 575 ..

1 MSPLPLRNP LTOEDRPFQE IIA

!IAA_SEQUENCE 1.0
P1:PH0858 - Maud protein - Paracoccus denitrificans (fragment)
C:Species: Paracoccus denitrificans
C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 08-Oct-1999
C:Accession: PH0858
R:Chistov, A.Y.; Boyd, J.; Mathews, F.S.; Lidstrom, M.E.
Biochem. Biophys. Res. Commun. 184, 1181-1189, 1992
A:Title: The genetic organization of the mau gene cluster of the facultative
autotroph Paracoccus denitrificans.
A:Reference number: PH0856; MUID:92272706; PMID:1350782
A:Accession: PH0858
A:Molecule type: DNA
A:Residues: 1-23 <CHI>
A:Cross-references: GB:M90098; NID:g150580; PIDN:AAA2557.1; PID:g150581
C:Genetics:
A:Gene: mauD

PH0858 Length: 23 December 22, 2002 19:21 Type: P Check: 807 ..

1 LEADRTGFAS LQQYMSRRK QAA

!IAA_SEQUENCE 1.0
P1:A61412 - methane monooxygenase (EC 1.14.13.25) protein Y - Methylococcus
capsulatus (fragment)
N:Alternate names: methane hydroxylase protein Y
C:Species: Methylococcus capsulatus
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Sep-1994
C:Accession: A61412
R:Rosenzweig, A.C.; Lipard, S.J.
Acc. Chem. Res. 27, 229-236, 1994
A:Title: Determining the structure of a hydroxylase enzyme that catalyzes the
conversion of methane to methanol in methanotrophic bacteria.
A:Reference number: A61412
A:Accession: A61412
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-27 <ROS>
C:Keywords: oxidoreductase

A61412 Length: 27 December 22, 2002 19:21 Type: P Check: 8072 ..

1 WIEDYASRID FKADRDQIVK AVLAGLK

!IAA_SEQUENCE 1.0
P1:A82288 - hypothetical protein VC0735 [imported] - Vibrio cholerae (strain
N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: A82288

```



```

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.;
Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill,
S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.;
Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P.; McDonald, L.;
Uttelback, T.; Fleischmann, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.;
Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.
A:Accession: AB2035; MUID:20406833; PMID:10952301
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-33 <HEI>
A:Cross-references: GB:AE004159; GB:AE003852; NID:9655167; PIDN:AAF93900.1;
GSPDB:GN00126; TIGR:VC0735
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0735
A:Map position: 1
A82288 Length: 33 December 22, 2002 19:21 Type: P Check: 3787 ..

1 MPVGSIPCS ISTALADRV VSRVWCFPL LVL

!!AA_SEQUENCE 1.0
P1:B82421 - hypothetical protein VCA0761 [Imported] - Vibrio cholerae (strain
N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: B82421
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.;
Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill,
S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.;
Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P.; McDonald, L.;
Uttelback, T.; Fleischmann, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.;
Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: B82421
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-32 <HEI>
A:Cross-references: GB:AE004404; GB:AE003853; NID:9658174; PIDN:AAF96659.1;
GSPDB:GN00127; TIGR:VCA0761
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0761
A:Map position: 2
B82421 Length: 32 December 22, 2002 19:21 Type: P Check: 1300 ..

1 MPDRLRFA ICSIDRKRQT NLFSSDLKQ WR

!!AA_SEQUENCE 1.0
F1:B39089 - hydorgenase (EC 1.18.99.1) 34k chain - Thiocapsa roseopersicina
(Fragment)
C:Species: Thiocapsa roseopersicina
C>Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 23-Jun-1993
C:Accession: B39089
R:Novacek, K.L.; Tsigyi, G.; Thanh, L.T.; Lakatos, S.; Kiss, Z.; Bagyinka, C.
J. Biol. Chem. 266, 947-951, 1991
A>Title: Structural rearrangements in active and inactive forms of hydorgenase
from Thiocapsa roseopersicina.
A:Reference number: A39089; MUID:91093297; PMID:1845998
A:Accession: B39089
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <KOY>
C:Keywords: oxidoreductase

B39089 Length: 20 December 22, 2002 19:21 Type: P Check: 6231 ..

1 XXEQADRPVS IGILPFGEXTG

!!AA_SEQUENCE 1.0
P1:H70954 - hypothetical protein RV3599c - Mycobacterium tuberculosis (strain
H37Rv)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: H70954
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garier, T.; Churcher, C.; Harris, D.;
Gordon, S.V.; Eigmeier, K.; Gas, S.; Barry III, C.E.; Tekala, F.; Badcock, K.;
Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.;
Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.;
Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail,
M.A.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.;
Squires, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70954
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-27 <COL>
A:Cross-references: GB:Z95557; GB:AL123456; NID:93242276; PIDN:CAB08945.1;
PID:9316847; PID:92113977
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: RV3599c
H70954 Length: 27 December 22, 2002 19:21 Type: P Check: 8577 ..

1 MPASSIGTGS PAADRDATH ERREVI

!!AA_SEQUENCE 1.0
F1:S14161 - probable dTDP-4-keio-6-deoxy-hexose 3,4-isomerase (EC 5.3.99.-) -
Saccharopolyspora erythraea (fragment)
C:Species: Saccharopolyspora erythraea
C>Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 21-Nov-1998
C:Accession: S14161
R:Caifrey, P.; Green, B.; Packman, L.C.; Rawlings, B.J.; Staunton, J.; Leadlay,
P.F.
Eur. J. Biochem. 195, 823-830, 1991
A>Title: An acyl-carrier-protein - thioesterase domain from the
6-deoxyerythronolide B synthase of Saccharopolyspora erythraea. High-level
production, purification and characterisation in Escherichia coli.
A:Reference number: S14091; MUID:91153324; PMID:1999200
A:Accession: S14161
A:Molecule type: protein
A:Residues: 1-20 <CAF>
C:Genetics:
A:Gene: eryCII
C:Function:
A:Description: involved in desosamine biosynthesis
C:Keywords: antibiotic biosynthesis; intramolecular oxidoreductase; isomerase
S14161 Length: 20 December 22, 2002 19:21 Type: P Check: 5957 ..

1 MTTTPRAGG ROLDMIRGLH

!!AA_SEQUENCE 1.0
F1:S66213 - glucose 1-dehydrogenase (EC 1.1.1.47) - Haloflex mediterranei
(Fragment)
C:Species: Haloflex mediterranei
C>Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 30-Oct-1998
C:Accession: S66213
R:Bonete, M.J.; Pire, C.; Ilorca, F.I.; Camacho, M.L.
FEBS Lett. 383, 227-229, 1996
A>Title: Glucose dehydrogenase from the halophilic Archaeon Haloflex
mediterranei: enzyme purification, characterisation and N-terminal sequence.

```

A:Reference number: S66213; MUID:96198607; PMID:8925901
 A:Accession: S66213
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-17 <BON>
 C:Keywords: oxidoreductase

S66213 Length: 17 December 22, 2002 19:21 Type: P Check: 1672 ..

1 MKAIIVKRG DRPVIE

!!AA_SEQUENCE 1.0
 FI:S23180 - carboxypeptidase - *Sulfolobus solfataricus*
 C:Species: *Sulfolobus solfataricus*
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
 C:Accession: S23180
 R:Colombo, S.; D'Auria, S.; Fusi, P.; Zecca, L.; Raia, C.A.; Tortora, P.
 Eur. J. Biochem. 206, 349-357, 1992
 A:Title: Purification and characterization of a thermostable carboxypeptidase from the extreme thermophilic archaeobacterium *Sulfolobus solfataricus*.
 A:Reference number: S23180; MUID:92283259; PMID:1597179
 A:Accession: S23180
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-31 <COL>

S23180 Length: 31 December 22, 2002 19:21 Type: P Check: 7956 ..

1 MDIVKLNKND VKIEDRIIQ IRRKIXENPP L

!!AA_SEQUENCE 1.0
 FI:S29326 - hypothetical protein 22, psba 5'-region - Japanese black pine chloroplast
 C:Species: chloroplast *Pinus thunbergiana* (Japanese black pine)
 C:Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 18-Aug-2000
 C:Accession: S29326; T07427
 R:Tsudzuki, J.; Nakashima, K.; Tsudzuki, T.; Hiratsuka, J.; Shibata, M.; Wakasugi, T.; Sugita, M.
 Mol. Gen. Genet. 232, 206-214, 1992
 A:Title: Chloroplast DNA of black pine retains a residual inverted repeat lacking rRNA genes: nucleotide sequences of trnQ, trnK, psba, trnI and trnH and the absence of rps16.
 A:Reference number: S20449; MUID:92212283; PMID:1557027

A:Accession: S29326
 A>Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-22 <TSU>
 A:Cross-references: EMBL:D11467; NID:9344007; PIDN:BAA02023.1; PID:9344011
 R:Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugita, M.
 Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994
 A:Title: Loss of all rnh genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*.
 A:Reference number: Z16030; MUID:95024047; PMID:7937893
 A:Accession: T07427
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-22 <MAK>
 A:Cross-references: EMBL:D17510; NID:9529643; PIDN:BAA04307.1; PID:91262588
 C:Genetics:
 A:Genome: chloroplast
 C:Keywords: chloroplast

S29326 Length: 22 December 22, 2002 19:21 Type: P Check: 9119 ..

1 MEYLTDRSI ECGIYLRKIE SI

!!AA_SEQUENCE 1.0
 FI:T10123 - probable catalase (EC 1.11.1.6) - cucumber (fragment)
 C:Species: *Cucumis sativus* (cucumber)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: T10123
 R:Toyama, T.; Teramoto, H.; Takeba, G.; Tsuji, H.
 Plant Cell Physiol. 36, 1349-1359, 1995

A:Title: Cytochrome c induces a rapid decrease in the levels of mRNAs for catalase, 3-hydroxy-3-methylglutaryl CoA reductase, lectin and other unidentified proteins in elicited *Arabidopsis thaliana* cells.

A:Reference number: Z16946; MUID:96104306; PMID:8564304
 A:Accession: T10123
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-23 <TOY>
 A:Cross-references: EMBL:D63385; NID:91199475; PIDN:BAA09701.1; PID:91199476
 A:Experimental source: seedling; cotyledons
 C:Keywords: oxidoreductase

T10123 Length: 23 December 22, 2002 19:21 Type: P Check: 1444 ..

1 NSQADRSVG QKLAPHNLVR PSI

!!AA_SEQUENCE 1.0
 FI:PS0212 - 29k protein 4228 - rice (strain Nihonbare) (fragment)
 C:Species: *Oryza sativa* (rice)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 11-Apr-1995
 C:Accession: PS0212
 R:Tsugita, A.; Miyatake, N.
 submitted to JIPID, April 1993
 A:Reference number: PS0208
 A:Accession: PS0212
 A:Molecule type: protein
 A:Residues: 1-15 <TSU>
 A:Experimental source: germ
 C:Comment: molecular weight 29k, pI 6.1.

PS0212 Length: 15 December 22, 2002 19:21 Type: P Check: 8963 ..

1 SPADRRDVG DRYAD

!!AA_SEQUENCE 1.0
 FI:B60698 - trichocyst protein 27 - *Paramecium tetraurelia* (fragment)
 C:Species: *Paramecium tetraurelia*
 C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 07-Dec-1999
 C:Accession: B60698
 R:Findall, S.H.; Devito, L.D.; Nelson, D.L.
 J. Cell Sci. 92, 441-447, 1989
 A:Title: Biochemical characterization of the proteins of *Paramecium* secretory granules.
 A:Reference number: A60698; MUID:90078398; PMID:2592449
 A:Accession: B60698
 A:Molecule type: protein
 A:Residues: 1-29 <TRN>
 C:Comment: This protein was purified from the extruded matrix (contents) of trichocysts (secretory granules).
 C:Genetics:
 A:Genetic code: SGC5
 C:Keywords: extracellular protein

B60698 Length: 29 December 22, 2002 19:21 Type: P Check: 2526 ..

1 DPIDRLSTL TDIEDRYVAE QKEDAKNQ

!!AA_SEQUENCE 1.0
 FI:S19614 - globin - polychaete (*Eudistylia vancouveri*) (fragment)
 N:Alternate names: chlorocruorin
 C:Species: *Eudistylia vancouveri*
 C:Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 05-Dec-1998
 C:Accession: S19614
 R:Qabar, A.N.; Stern, M.S.; Walz, D.A.; Chiu, J.T.; Timkovich, R.; Wall, J.S.; Kapp, O.H.; Vinogradov, S.N.
 J. Mol. Biol. 222, 1109-1129, 1991
 A:Title: Hierarchy of globin complexes. The quaternary structure of the extracellular chlorocruorin of *Eudistylia vancouveri*.
 A:Reference number: S19532; MUID:92106333; PMID:1762147
 A:Accession: S19614
 A:Molecule type: protein
 A:Residues: 1-17 <QAB>
 A:Experimental source: plume

C;Complex: dodecamers, each consisting of a trimer of tetramers of globin chains; dodecamers are linked into bilayer structure by Ca(2+) and heme-deficient chimeric globin chains
C;Keywords: calcium; dodecamer; heme; homotetramer; oxygen carrier
SI19614 Length: 17 December 22, 2002 19:21 Type: P Check: 1820 ..

1 GMRXXSMEDR KTVLADW

!IAA_SEQUENCE 1.0
FI:A60741 - insulin-like growth factor-binding protein He39L - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 30-Sep-1993
C;Accession: A60741
R;Forbes, B.; Ballard, F.D.; Wallace, J.C.
J. Endocrinol. 126, 497-506, 1990
A;Title: An insulin-like growth factor-binding protein purified from medium conditioned by a human lung fibroblast cell line (He139L) has a novel N-terminal sequence.
A;Reference number: A60741; MUID:91011238; PMID:1698907
A;Accession: A60741
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-25 <FOR>

A60741 Length: 25 December 22, 2002 19:21 Type: P Check: 4352 ..

1 LAPPGGQGVQ AGPGGXVEE EDRGS

!IAA_SEQUENCE 1.0
FI:B28457 - proteoglycan II, bone - human (tentative sequence) (fragment)
C;Species: Homo sapiens (man)
C;Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 31-Mar-2000
C;Accession: B28457
R;Fisher, L.W.; Hawkins, G.R.; Tuross, N.; Termine, J.D.
J. Biol. Chem. 262, 9702-9708, 1987
A;Title: Purification and partial characterization of small proteoglycans I and II, bone sialoproteins I and II, and osteonectin from the mineral compartment of developing human bone.
A;Reference number: A92656; MUID:87250639; PMID:3597437
A;Accession: B28457
A;Molecule type: protein
A;Residues: 1-19 <FTS>

B28457 Length: 19 December 22, 2002 19:21 Type: P Check: 4360 ..

1 DEAXGIAPEV PDDRFEPES

!IAA_SEQUENCE 1.0
FI:C54037 - splicing regulatory protein SWAP homolog (alternatively spliced, clone pFL2) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 28-Feb-1997
C;Accession: C54037
R;Denhez, F.; Laliyalis, R.
J. Biol. Chem. 269, 16170-16179, 1994
A;Title: Conservation of regulated alternative splicing and identification of functional domains in vertebrate homologs to the Drosophila splicing regulator, suppressor-of-white-apricot
A;Reference number: A54037; MUID:94266805; PMID:8206918
A;Accession: C54037
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-29 <DEN>
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBIP:149928)

C54037 Length: 29 December 22, 2002 19:21 Type: P Check: 1565 ..

1 POLTOELEEA KOAKOLEDR LAAAREKL

!IAA_SEQUENCE 1.0
FI:A48845 - sterol regulatory element 1 binding protein (alternatively spliced,

clone pCY22) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C;Accession: A48845
R;Yokoyama, C.; Wang, X.; Briggs, M.R.; Admon, A.; Wu, J.; Hua, X.; Goldstein, J.L.; Brown, M.S.
Cell 75, 187-197, 1993
A;Title: SREBP-1, a basic-helix-loop-helix-leucine zipper protein that controls transcription of the low density lipoprotein receptor gene.
A;Reference number: A48845; MUID:94006541; PMID:8402897
A;Accession: A48845
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-23 <YOK>
A;Cross-references: GB:S66168; MUID:9432596; PID:MAB26522.1; PID:9432597
A;Experimental source: HeLa S3 cells
A;Note: sequence extracted from NCBI backbone (NCBIN:138566, NCBIP:138570)

A48845 Length: 23 December 22, 2002 19:21 Type: P Check: 9975 ..

1 KGRANGTDAP RAGADRGAMD CTF

!IAA_SEQUENCE 1.0
FI:S22228 - vitronectin - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 22-Nov-1993 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
C;Accession: S22228
R;Nakashima, N.; Miyazaki, K.; Ishikawa, M.; Yatohgo, T.; Ogawa, H.; Uchibori, H.; Matsumoto, I.; Seno, N.; Hayashi, M.
Biochim. Biophys. Acta 1120, 1-10, 1992
A;Title: Vitronectin diversity in evolution but uniformity in ligand binding and size of the core polypeptide
A;Reference number: S21768; MUID:92207982; PMID:1372829
A;Accession: S22228
A;Molecule type: protein
A;Residues: 1-25 <NAK>

S22228 Length: 25 December 22, 2002 19:21 Type: P Check: 4810 ..

1 AOESXKGRVT EGFNADRKQO QDELY

!IAA_SEQUENCE 1.0
FI:A26393 - annexin 36K chain - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 02-Jun-1988 #sequence_revision 02-Jun-1988 #text_change 18-Jun-1993
C;Accession: A26393
R;Geisow, M.J.; Fritsche, U.; Hexham, J.M.; Dash, B.; Johnson, T.
Nature 320, 636-638, 1986
A;Title: A consensus amino-acid sequence repeat in Torpedo and mammalian Ca(2+)-dependent membrane-binding proteins
A;Reference number: A93379; MUID:86203621; PMID:2422556
A;Accession: A26393
A;Molecule type: protein
A;Residues: 1-16 <GEI>

A26393 Length: 16 December 22, 2002 19:21 Type: P Check: 7 ..

1 MGLGTDDEDR LITETLL

!IAA_SEQUENCE 1.0
FI:S68260 - hypothetical protein gadd7.1 - long-tailed hamster
C;Species: Critetulus longicaudatus (long-tailed hamster)
C;Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C;Accession: S68260
R;Hollander, M.C.; Alamo, I.; Fornace Jr., A.J.
Nucleic Acids Res. 24, 1589-1593, 1996
A;Title: A novel DNA damage-inducible transcript, gadd7, inhibits cell growth, but lacks a protein product.
A;Reference number: S68260; MUID:96211359; PMID:8649973
A;Accession: S68260
A;Status: preliminary
A;Molecule type: mRNA

```
A:Residues: 1-38 <HOL>
A:Cross-references: EMBL:L40430

S68260 Length: 38 December 22, 2002 19:21 Type: P Check: 6375 ..

1 MPFSFGSSK QLQRNAQMED RGNHPSEEFI TGDNLKLT

!!AA_SEQUENCE 1.0
P1:A35678 - hypothetical protein (proenkephalin 5' region) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 05-Nov-1999
C:Accession: A35678
R:Kilpatrick, D.L.; Zinn, S.A.; Fitzgerald, M.; Higuchi, H.; Sabol, S.L.;
Meyershard, J.
Mol. Cell. Biol. 10, 3717-3726, 1990
A:Title: Transcription of the rat and mouse proenkephalin genes is initiated at
distinct sites in spermatogenic and somatic cells.
A:Reference number: A35678; MUID:90287163; PMID:2355920
A:Accession: A35678
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-18 <KLT>
A:Cross-references: GB:M55181; NID:9201032; PIDN:AAA40127.1; PID:9201033

A35678 Length: 18 December 22, 2002 19:21 Type: P Check: 3193 ..

1 MSSGKQDSPW EDRIIPGR

!!AA_SEQUENCE 1.0
F1:PH0780 - T-cell receptor alpha chain (C11) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PH0780
R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A:Title: T cell receptor genes in a series of class I major histocompatibility
complex-restricted cytotoxic T lymphocyte clones specific for a Plasmodium
berghel nonapeptide: implications for T cell allelic exclusion and
antigen-specific repertoire.
A:Reference number: PH0746; MUID:92078846; PMID:1836010
A:Accession: PH0780
A:Molecule type: mRNA
A:Residues: 1-15 <CAS>
A:Cross-references: EMBL:X60879
A:Experimental source: T lymphocyte
C:Keywords: T-cell receptor

PH0780 Length: 15 December 22, 2002 19:21 Type: P Check: 8887 ..

1 CALSETGAD RLTFG

!!AA_SEQUENCE 1.0
F1:PH0794 - T-cell receptor alpha chain (K1 V-alpha-3-PHDS58) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PH0794
R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A:Title: T cell receptor genes in a series of class I major histocompatibility
complex-restricted cytotoxic T lymphocyte clones specific for a Plasmodium
berghel nonapeptide: implications for T cell allelic exclusion and
antigen-specific repertoire.
A:Reference number: PH0746; MUID:92078846; PMID:1836010
A:Accession: PH0794
A:Molecule type: mRNA
A:Residues: 1-17 <CAS>
A:Cross-references: EMBL:X60899
A:Experimental source: T lymphocyte
C:Keywords: T-cell receptor

PH0794 Length: 17 December 22, 2002 19:21 Type: P Check: 1531 ..

1 CAVSMNEYRG ADRLTFG

A:Residues: 1-10 <NAK>
A:Cross-references: EMBL:X02859
A:Experimental source: T lymphocyte
C:Keywords: T-cell receptor

S03505 Length: 20 December 22, 2002 19:21 Type: P Check: 6013 ..

1 NTEGADRLTF GKGTQLTIQP

!!AA_SEQUENCE 1.0
F1:PT0212 - T-cell receptor alpha chain V-J region (4-1-E.2) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 30-May-1997
C:Accession: PT0212
R:Nakano, N.; Kikutani, H.; Nishimoto, H.; Kishimoto, T.
J. Exp. Med. 173, 1091-1097, 1991
A:Title: T cell receptor V gene usage of islet beta cell-reactive T cells is
not restricted in non-obese diabetic mice.
A:Reference number: PT0209; MUID:91217621; PMID:1902501
A:Accession: PT0212
A:Molecule type: mRNA
A:Residues: 1-10 <NAK>
C:Keywords: T-cell receptor

PT0212 Length: 10 December 22, 2002 19:21 Type: P Check: 3993 ..

1 CAVAGGADRL

!!AA_SEQUENCE 1.0
F1:PT0547 - T-cell receptor beta chain V-D-J region (126-1A1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
```

```

C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0547
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N
regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0547
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-8 <FEF>
A:Experimental source: day 18 fetal thymus, strain BALB/c
C:Keywords: T-cell receptor

PT0547 Length: 8 December 22, 2002 19:21 Type: P Check: 2627 ..

1 ASSDADRG

!!AA_SEQUENCE 1.0
F1:PT0576 - T-cell receptor beta chain V-D-J region (140-14L) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0676
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N
regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0676
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-7 <FEF>
A:Experimental source: day 18 fetal thymus, strain BALB/c
C:Keywords: T-cell receptor

PT0676 Length: 7 December 22, 2002 19:21 Type: P Check: 2049 ..

1 ASGEDRG

!!AA_SEQUENCE 1.0
F1:PT0576 - T-cell receptor beta chain V-D-J region (141-16) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0576
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N
regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0576
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-7 <FEF>
A:Experimental source: day 19 fetal thymus, strain BALB/c
C:Keywords: T-cell receptor

PT0576 Length: 7 December 22, 2002 19:21 Type: P Check: 2172 ..

1 ASSDRT

!!AA_SEQUENCE 1.0
F1:PT0366 - T-cell receptor beta chain V-J region (6R2) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 30-May-1997
C:Accession: PT0366
R:Lehmann, P.V.; Drexler, K.; Tary-Lehmann, M.; Falcioni, F.; Huttenbach, U.;
Nagy, Z.A.
J. Exp. Med. 173, 333-341, 1991
A:Title: Graft-versus-host resistance induced by class II major
histocompatibility complex-specific T cell clones.
A:Reference number: PT0360; MUID:91108330; PMID:1824856
A:Accession: PT0366
A:Molecule type: mRNA

```

```

A:Residues: 1-28 <LEH>
C:Keywords: T-cell receptor

PT0366 Length: 28 December 22, 2002 19:21 Type: P Check: 480 ..

1 LYFCASSEDR NNOLFLERGLDFSVLED

!!AA_SEQUENCE 1.0
F1:A46592 - lactase-phlorizin hydrolase, 200K isoform - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 01-Nov-1996
C:Accession: A46592
R:Dudley, M.A.; Hachey, D.L.; Quaroni, A.; Hutchens, T.W.; Nichols, B.L.;
Rosenberger, J.; Perkinson, J.S.; Cook, G.; Reeds, P.J.
J. Biol. Chem. 268, 13609-13616, 1993
A:Title: In vivo sucrose-isomaltase and lactase-phlorizin hydrolase turnover in
the fed adult rat.
A:Reference number: A46592; MUID:93293888; PMID:8514793
A:Accession: A46592
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-17 <DUU>
A>Note: sequence extracted from NCBI backbone (NCBIP:134559)
C:Keywords: carbohydrate digestion; intestine

A46592 Length: 17 December 22, 2002 19:21 Type: P Check: 1363 ..

1 DMEDNRFTAA GPLTNDL

!!AA_SEQUENCE 1.0
F1:S78414 - ribosomal protein RL25, mitochondrial [validated] - rat (tentative
sequence) (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 25-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 21-Jul-2000
C:Accession: S78414
R:Goldschmidt-Reisin, S.; Graack, H.R.
submitted to the Protein Sequence Database, February 1998
A:Reference number: S78411
A:Accession: S78414
A:Molecule type: protein
A:Residues: 1-26 <GOL>
A>Note: the protein is designated as mitochondrial ribosomal protein L25
C:Keywords: mitochondrion; protein biosynthesis; ribosome

S78414 Length: 26 December 22, 2002 19:21 Type: P Check: 6086 ..

1 ISRRXEKRNK IYPDQLDGE DRRDAE

!!AA_SEQUENCE 1.0
F1:S22227 - vitronectin - guinea pig (fragment)
C:Species: Cavia porcellus (guinea pig)
C>Date: 22-Nov-1993 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
C:Accession: S22227
R:Nakashima, N.; Miyazaki, K.; Ishikawa, M.; Yatohgo, T.; Ogawa, H.; Uchibori,
H.; Matsumoto, I.; Seno, N.; Hayashi, M.
Biochim. Biophys. Acta 1120, 1-10, 1992
A:Title: Vitronectin diversity in evolution but uniformity in ligand binding
and size of the core polypeptide.
A:Reference number: S21768; MUID:92207982; PMID:1372829
A:Accession: S22227
A:Molecule type: protein
A:Residues: 1-25 <NAK>

S22227 Length: 25 December 22, 2002 19:21 Type: P Check: 5585 ..

1 XXESXKGRXT EGFNADRKXQ XXELX

!!AA_SEQUENCE 1.0
F1:A54226 - light-harvesting protein B-830 alpha-1 chain - Chromatium
purpuratum (fragment)
C:Species: Chromatium purpuratum
C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 05-Jan-1996
C:Accession: A54226

```

R.Kerfeld, C.A.; Yeates, T.O.; Thorner, J.P.
 Biochemistry 33, 2178-2184, 1994
 A:Title: Purification and characterization of the peripheral antenna of the purple-sulfur bacterium Chromatium purpuratum: evidence of an unusual pigment-protein composition.
 A:Reference number: A54226; MUID:94162224; PMID:8117674
 A:Accession: A54226
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-26 <KER>
 C:Keywords: antenna complex; light-harvesting polypeptide

A54226 Length: 26 December 22, 2002 19:21 Type: P Check: 6170 ..

1 MKVPMMADE NAKLNPPEDD RKKFFV

!IAA_SEQUENCE 1.0
 F1:B54226 - light-harvesting protein B-830 alpha-2 chain - Chromatium purpuratum (fragment)
 C:Species: Chromatium purpuratum
 C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 05-Jan-1996
 C:Accession: B54226
 R:Kerfeld, C.A.; Yeates, T.O.; Thorner, J.P.
 Biochemistry 33, 2178-2184, 1994
 A:Title: Purification and characterization of the peripheral antenna of the purple-sulfur bacterium Chromatium purpuratum: evidence of an unusual pigment-protein composition.
 A:Reference number: A54226; MUID:94162224; PMID:8117674
 A:Accession: B54226
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-23 <KER>
 C:Keywords: antenna complex; light-harvesting polypeptide

B54226 Length: 23 December 22, 2002 19:21 Type: P Check: 634 ..

1 MQVPMIADK NAKLNHPEDD RKR

!IAA_SEQUENCE 1.0
 F1:T46593 - phytoene dehydrogenase [imported] - Mycobacterium marinum (fragment)
 C:Species: Mycobacterium marinum
 C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 21-Jul-2000
 C:Accession: T46593
 R:Ramakrishnan, L.; Tran, H.T.; Federspiel, N.A.; Falkow, S.
 J. Bacteriol. 179, 5862-5868, 1997
 A:Title: A crtB homolog essential for photochromogenicity in Mycobacterium marinum: isolation, characterization, and gene disruption via homologous recombination.
 A:Reference number: T23096; MUID:97440138; PMID:9294446
 A:Accession: T46593
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-38 <RAM>
 A:Cross-references: EMBL:U92075; NID:g1928930; PIDN:AAB71427.1; PID:g1928931
 A:Experimental source: strain M
 C:Genetics:
 A:Gene: crtI

T46593 Length: 38 December 22, 2002 19:21 Type: P Check: 6757 ..

1 VPGVGPTTL ISGRLAARI TGNTTRSRH LDLKAQLS

!IAA_SEQUENCE 1.0
 P1:H85575 - hypothetical protein Z0899 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85575
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, G.F.; Evans, P.S.; Gregor, J.; Kirkpatrick, H.A.; Postal, G.; Hackett, J.; Klink, S.; Boutin, A.; Shao, Y.; Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, J.; Anantharaman, T.S.; Llin, J.; Yen, G.; Schwartz, D.C.; Welch, R.A.; Blattner, F.R.

Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: H85575
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-13 <STO>
 A:Cross-references: GB:AE005174; NID:g12513665; PIDN:AAG55068.1; GSPDB:GN00145; UWGP:Z0899
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: Z0899

H85575 Length: 13 December 22, 2002 19:21 Type: P Check: 6940 ..

1 MSTDRKPYML LFH

!IAA_SEQUENCE 1.0
 F1:PA0041 - plastocyanin-plastocyanin reductase (EC 1.10.99.1) - Arabidopsis thaliana (fragment)
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 30-Jun-1992 #sequence_revision 06-Jan-1995 #text_change 03-Jun-2002
 C:Accession: PA0041
 R:Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.
 submitted to JPIID, July 1994
 A:Description: Separation and characterization of Arabidopsis proteins by two-dimensional gel electrophoresis.
 A:Reference number: PA0001
 A:Accession: PA0041
 A:Molecule type: protein
 A:Residues: 1-15 <KAM>
 A:Experimental source: leaf
 C:Keywords: oxidoreductase

PA0041 Length: 15 December 22, 2002 19:21 Type: P Check: 9117 ..

1 ASSIPADRVV DMEKR

```

11AA_SEQUENCE 1.0
ID AL20_CARMA STANDARD; PRT; 27 AA.
AC P81823;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinus maenas (Common shore crab).
OS Carcinus maenas (Common shore crab).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC Brachyura; Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_Taxid=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=96121193; PubMed=9461295;
RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jares P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
CC Neuropeptide; Amidation; Multigene family.
FT MOD_RES 27 AA; 3152 MW; 597A6901965FE967 CRC64;
SQ SEQUENCE 27 AA; 3152 MW; 597A6901965FE967 CRC64;

AL20_CARMA Length: 27 December 22, 2002 19:30 Type: P Check: 9046

1 GYEDDEDHP FYALGLKRP RYSPGL

11AA_SEQUENCE 1.0
ID AMD1_CHICK STANDARD; PRT; 26 AA.
AC P81073;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE AMP deaminase 1 (EC 3.5.4.6) (Myoadenylate deaminase) (AMP deaminase
DE isoform M) (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_Taxid=9031;
RN [1]
RP SEQUENCE.
RC TISSUE=Breast;
RX MEDLINE=97269365; PubMed=9114497;
RA Chilson O.P., Kelly-Chilson A.E., Siegel N.R.;
RT "AMP-deaminases from chicken and rabbit muscle: partial primary
RT sequences of homologous 17-kDa CNBR fragments: autorecognition by
RT rabbit anti-(chicken AMP).";
RL Comp. Biochem. Physiol. 116B:371-377(1997).
CC -1- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY
CC METABOLISM.
CC -1- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).
CC -1- PATHWAY: PURINE NUCLEOTIDE CYCLE.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.
DR InterPro: IPR001365; A/AMP deaminase.
DR PROSITE: PS00485; A/AMP deaminase.
KW PROSITE; PS00485; A/AMP deaminase.
KW Hydroxylase; Nucleotide metabolism; Multigene family.
FT NON_TER 1
FT NON_TER 26
SQ SEQUENCE 26 AA; 3195 MW; B03E296D63B6E75 CRC64;

AMD1_CHICK Length: 26 December 22, 2002 19:30 Type: P Check: 7006

1 MNQKHLRFI KSYRVADNR VYVDAR

11AA_SEQUENCE 1.0
ID AMD1_RABIT STANDARD; PRT; 26 AA.
AC P81072;

```

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE AMP deaminase 1 (EC 3.5.4.6) (Myoadenylate deaminase) (AMP deaminase
DE isoform M) (Fragment).
GN AMPD1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE=97269365; PubMed=9114497;
RA Chilson O.P., Kelly-Chilson A.E., Siegel N.R.;
RT "AMP-deaminases from chicken and rabbit muscle: partial primary
RT sequences of homologous 17-kDa CNBR fragments: autorecognition by
RT rabbit anti-(chicken AMP).";
RL Comp. Biochem. Physiol. 116B:371-377(1997).
CC -1- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY
CC METABOLISM.
CC -1- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).
CC -1- PATHWAY: PURINE NUCLEOTIDE CYCLE.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP
CC DEAMINASE 1 IS THE PREDOMINANT FORM IN SMOOTH MUSCLE, NON-MUSCLE TISSUE,
CC EMBRYONIC MUSCLE AND UNDIFFERENTIATED MYOBLASTS; AMP DEAMINASE 3
CC IS FOUND IN ERYTHROCYTES.
CC -1- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.
DR InterPro: IPR001365; A/AMP deaminase.
DR PROSITE: PS00485; A/AMP deaminase.
KW Hydroxylase; Nucleotide metabolism; Multigene family.
FT NON_TER 1
FT NON_TER 26
SQ SEQUENCE 26 AA; 3169 MW; B022467BACB6E75 CRC64;

AMD1_RABIT Length: 26 December 22, 2002 19:30 Type: P Check: 7826

1 MNQKHLRFI KSYRVADNR VYVSTR

11AA_SEQUENCE 1.0
ID DNIY_BPD10 STANDARD; PRT; 22 AA.
AC Q38199;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA-invertase (Fragment).
GN GIN.
OS Bacteriophage D108.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC Mu-like viruses.
OX NCBI_Taxid=10671;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87316928; PubMed=2957646;
RA Szatmari G.B., Lapointe M., Dubow M.S.;
RT "The right end of transposable bacteriophage D108 contains a 520 base
RT pair protein-encoding sequence not present in bacteriophage Mu.";
RL Nucleic Acids Res. 15:6691-6703(1987).
CC -1- FUNCTION: THIS PROTEIN CATALYZES THE INVERSION OF A 3000-BP
CC SEGMENT OF PHAGE DNA. THE INVERSION RESULTS IN A MODIFICATION OF
CC THE 3' END OF THE TAIL FIBER GENE AND ALTERS THE HOST SPECIFICITY.
CC -1- SIMILARITY: BELONGS TO THE SITE-SPECIFIC RECOMBINASE RESOLVASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: X05926; CAA29365.1;
DR InterPro: IPR001822; RECOMBINASE.
DR PROSITE: PS00397; RECOMBINASE_1; PARTIAL.
DR PROSITE: PS00398; RECOMBINASE_2; PARTIAL.
KM DNA recombination; DNA integration; DNA-binding; DNA invertase.
FT NON_TER 1 1
FT MOD_RES 1 1 H-T-H MOTIF (PROBABLY).
FT MOD_RES 21 21 Sulfation.
SQ SEQUENCE 22 AA; 2748 MW; 09E607032767C38 CRC64;
DNV_BPD10 Length: 22 December 22, 2002 19:30 Type: P Check: 9004 ..

1 YKKHAKRTH IENDRINOI DR
!!A_SEQUENCE 1.0
ID FIBB_ANTAM STANDARD; PRT; 21 AA.
AC P14465;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [contains: fibrinopeptide B] (Fragment).
GN FGB.
OS Antilocapra americana (Pronghorn).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Antilocapridae; Antilocapra.
OX NCBI_TaxID=9891;
RN [1]
RP SEQUENCE.
RA Moss G.A., Doolittle R.F.;
RT "amino acid sequence studies on artiodactyl fibrinopeptides.";
RL Arch. Biochem. Biophys. 122:674-684(1967).
CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
CC InterPro: IPR002181; Fibrinogen.C.
DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KM Blood coagulation; Plasma; Sulfation.
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 6 6 SULFATION.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2585 MW; FCE6183BE1F31627 CRC64;
FIBB_ANTAM Length: 21 December 22, 2002 19:30 Type: P Check: 7016 ..

1 QPSYDYDEER DDBAKRLIDA R
!!A_SEQUENCE 1.0
ID FIBB_BISBO STANDARD; PRT; 21 AA.
AC P14466;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [contains: fibrinopeptide B] (Fragment).
GN FGB.
OS Bison bonasus (European bison).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bison.
OX NCBI_TaxID=9902;
RN [1]
RP SEQUENCE.
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET

```

```

CC AGGREGATION.
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
CC InterPro: IPR002181; Fibrinogen.C.
DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KM Blood coagulation; Plasma; Sulfation.
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
FT MOD_RES 6 6 SULFATION.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2366 MW; 09E675AF19E5363D CRC64;
FIBB_BISBO Length: 21 December 22, 2002 19:30 Type: P Check: 7213 ..

1 EPPDYDEGE DDBPKVLGA R
!!A_SEQUENCE 1.0
ID FIBB_BUBBU STANDARD; PRT; 21 AA.
AC P14467;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [contains: fibrinopeptide B] (Fragment).
GN FGB.
OS Bubalus bubalis (Domestic water buffalo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bubalus.
OX NCBI_TaxID=89462;
RN [1]
RP SEQUENCE.
RA Moss G.A., Doolittle R.F.;
RT "amino acid sequence studies on artiodactyl fibrinopeptides.";
RL Arch. Biochem. Biophys. 122:674-684(1967).
CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
CC InterPro: IPR002181; Fibrinogen.C.
DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KM Blood coagulation; Plasma; Sulfation.
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
FT MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 21 21 SULFATION.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2379 MW; 09E675BE4729163D CRC64;
FIBB_BUBBU Length: 21 December 22, 2002 19:30 Type: P Check: 7185 ..

1 QPPTDYDEGE DDBPKVLGA R
!!A_SEQUENCE 1.0
ID FIBB_CEREL STANDARD; PRT; 21 AA.
AC P14468;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```



```

DE  Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN  FGB.
OS  Cervus elaphus (Red deer), and
OC  Cervus elaphus nelsoni (American elk).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC  Cervidae; Cervinae; Cervus.
OX  NCBI_TaxID=9860, 9864;
RN  [1]
RP  SEQUENCE.
RC  SPECIES=C. elaphus;
RA  Blomback B., Blomback M., Grondahl N.J.;
RT  "Studies on fibrinopeptides from mammals.";
RL  Acta Chem. Scand. 19:1789-1791(1965).
RN  [2]
RP  SEQUENCE.
RC  SPECIES=C. e. nelsoni;
RA  Moss G.A., Doolittle R.F.;
RT  "Amino acid sequence studies on aridodactyl fibrinopeptides.";
RL  Arch. Biochem. Biophys. 122:674-684(1967).
CC  -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC  POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC  AGGREGATION.
CC  -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC  (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC  -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC  THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC  CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC  RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
CC  Interpro: IPR002181; Fibrinogen.C.
DR  PROSITE: PS00514; FIBRIN_AG_C_DOMAIN: PARTIAL.
KW  Blood coagulation; Plasma; Sulfation.
FT  PEPTIDE 1 21 FIBRINOPEPTIDE B.
FT  MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT  NON_TER 6 6 SULFATION.
FT  NON_TER 21 21
SQ  SEQUENCE 21 AA: 2558 MW: FCEE745D98931627 CRC64;

FIBB_CEREL Length: 21 December 22, 2002 19:30 Type: P Check: 6821 ..

1 OHSTDYDEE EDRAKLHLDA R

!!AA_SEQUENCE 1.0
ID  FIBB_FELCA STANDARD: PRT: 20 AA.
AC  P14469;
DT  01-JAN-1990 (Rel. 13, Created)
DT  01-JAN-1990 (Rel. 13, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN  FGB.
OS  Felis silvestris catus (Cat).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX  NCBI_TaxID=9685;
RN  [1]
RP  SEQUENCE.
RA  Blomback B., Blomback M., Grondahl N.J.;
RT  "Studies on fibrinopeptides from mammals.";
RL  Acta Chem. Scand. 19:1789-1791(1965).
CC  -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC  POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC  AGGREGATION.
CC  -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC  (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC  -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC  THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC  CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC  RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
CC  Interpro: IPR002181; Fibrinogen.C.
DR  PROSITE: PS00514; FIBRIN_AG_C_DOMAIN: PARTIAL.
KW  Blood coagulation; Plasma.
FT  PEPTIDE 1 20 FIBRINOPEPTIDE B.
FT  NON_TER 20 20

```

```

SQ  SEQUENCE 20 AA: 2328 MW: A829E393B8F627D0 CRC64;

FIBB_FELCA Length: 20 December 22, 2002 19:30 Type: P Check: 5816 ..

1 IIDYDEGE DRDVGVDAR

!!AA_SEQUENCE 1.0
ID  FIBB_LAMGL STANDARD: PRT: 19 AA.
AC  P14473;
DT  01-JAN-1990 (Rel. 13, Created)
DT  01-JAN-1990 (Rel. 13, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN  FGB.
OS  Lama glama (Llama),
OS  Lama vicugna (Vicugna) (Vicugna vicugna), and
OS  Camelus dromedarius (Dromedary) (Arabian camel).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
OX  NCBI_TaxID=9844, 9843, 9838;
RN  [1]
RP  SEQUENCE.
RC  SPECIES=L. glama;
RA  Blomback B., Blomback M., Grondahl N.J.;
RT  "Studies on fibrinopeptides from mammals.";
RL  Acta Chem. Scand. 19:1789-1791(1965).
RN  [2]
RP  SEQUENCE.
RC  SPECIES=C. dromedarius;
RX  MEDLINE=67209145; Pubmed=6033721;
RA  Doolittle R.F., Schubert D., Schwartz S.A.;
RT  "Amino acid sequence studies on aridodactyl fibrinopeptides. I.
RT  Dromedary camel, mule deer, and cape buffalo.";
RL  Arch. Biochem. Biophys. 118:456-467(1967).
RN  [3]
RP  SEQUENCE.
RC  SPECIES=L. vicugna;
RA  Moss G.A., Doolittle R.F.;
RT  "Amino acid sequence studies on aridodactyl fibrinopeptides.";
RL  Arch. Biochem. Biophys. 122:674-684(1967).
CC  -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC  POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC  AGGREGATION.
CC  -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC  (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC  -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC  THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC  CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC  RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
CC  Interpro: IPR002181; Fibrinogen.C.
DR  PROSITE: PS00514; FIBRIN_AG_C_DOMAIN: PARTIAL.
KW  Blood coagulation; Plasma; Sulfation.
FT  PEPTIDE 1 19 FIBRINOPEPTIDE B.
FT  MOD_RES 4 4 SULFATION.
FT  NON_TER 19 19
SQ  SEQUENCE 19 AA: 2295 MW: E7EE6B6100568638 CRC64;

FIBB_LAMGL Length: 19 December 22, 2002 19:30 Type: P Check: 4317 ..

1 ATDYDEED RVRVLDAR

!!AA_SEQUENCE 1.0
ID  FIBB_MUNMU STANDARD: PRT: 21 AA.
AC  P14475;
DT  01-JAN-1990 (Rel. 13, Created)
DT  01-FEB-1994 (Rel. 28, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN  FGB.
OS  Muntiacus muntjak (Muntjak).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC  Cervidae; Muntiacinae; Muntiacus.

```

OX NCBI_TaxID=9888;
 RN [1]
 RP SEQUENCE.
 RA "Moss G.A., Doolittle R.F.;"
 RL "Amino acid sequence studies on arthropod fibrinopeptides.";
 Arch. Biochem. Biophys. 122:674-684(1967).
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 DR Interpro: IPR002181; Fibrinogen_C.
 KW PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KM Blood coagulation; Plasma; Sulfation.
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 6 6 SULFATION.
 FT NON_TER 21 21
 SQ SEQUENCE 21 AA; 2514 MW; FCFE5188F0C1627 CRC64;
 FIBB_MUNMG Length: 21 December 22, 2002 19:30 Type: P Check: 6963 ..

1 QHLDYDEV DRAKLHDA R
 ID FIBB_ODOHE STANDARD; PRT; 21 AA.
 AC P14476;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [contains: Fibrinopeptide B] (Fragment).
 GN FGB.
 OS Odocoileus hemionus (Mule deer) (Black-tailed deer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
 OC Cervidae; Odocoileinae; Odocoileus.
 OX NCBI_TaxID=9872;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE-67209145; PubMed-6033721;
 RA Doolittle R.F., Schubert D., Schwartz S.A.;"
 RL "Amino acid sequence studies on arthropod fibrinopeptides. I.
 RT Dromedary camel, mule deer, and cape buffalo.";
 Arch. Biochem. Biophys. 118:456-467(1967).
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 DR Interpro: IPR002181; Fibrinogen_C.
 KW PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KM Blood coagulation; Plasma; Sulfation.
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 6 6 SULFATION.
 FT NON_TER 21 21
 SQ SEQUENCE 21 AA; 2496 MW; FCF562C51A0C1627 CRC64;
 FIBB_ODOHE Length: 21 December 22, 2002 19:30 Type: P Check: 6856 ..

DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [contains: Fibrinopeptide B] (Fragment).
 GN FGB.
 OS Rangifer tarandus (Reindeer) (Caribou).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
 OC Cervidae; Odocoileinae; Rangifer.
 OX NCBI_TaxID=9870;
 RN [1]
 RP SEQUENCE.
 RA Blomback B., Blomback M., Gronlund N.J.;"
 RL "Studies on fibrinopeptides from mammals.";
 Acta Chem. Scand. 19:1789-1791(1965).
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 DR Interpro: IPR002181; Fibrinogen_C.
 KW PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KM Blood coagulation; Plasma; Sulfation.
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 6 6 SULFATION.
 FT NON_TER 21 21
 SQ SEQUENCE 21 AA; 2510 MW; FCF562C45F0C1627 CRC64;
 FIBB_RANTA Length: 21 December 22, 2002 19:30 Type: P Check: 6866 ..

1 QHLDYDEV DRAKLHDA R
 ID FIBB_SYMCA STANDARD; PRT; 21 AA.
 AC P14481;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [contains: Fibrinopeptide B] (Fragment).
 GN FGB.
 OS Syncerus caffer (Cape buffalo).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Syncerus.
 OX NCBI_TaxID=9970;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE-67209145; PubMed-6033721;
 RA Doolittle R.F., Schubert D., Schwartz S.A.;"
 RL "Amino acid sequence studies on arthropod fibrinopeptides. I.
 RT Dromedary camel, mule deer, and cape buffalo.";
 Arch. Biochem. Biophys. 118:456-467(1967).
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 DR Interpro: IPR002181; Fibrinogen_C.
 KW PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KM Blood coagulation; Plasma; Sulfation.
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 6 6 SULFATION.
 FT NON_TER 21 21

SQ SEQUENCE 21 AA; 2353 MW; 09E75AF03B863D CRC64;
 FTB_SYNCA Length: 21 December 22, 2002 19:30 Type: P Check: 7177 ..
 1 QPPTDYDEGE DMRPRSGIGA R
 11AA_SEQUENCE 1.0 STANDARD: PRT: 36 AA.
 ID GLU1_0RENI
 AC P81026;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Glucagon 1.
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthomorphi; Acanthopterygii; Perciformes; Labroidae; Cichlidae; Oreochromis.
 NC NCB1_TaxID=8128;
 RX MEDLINE=9384941; PubMed=7656183;
 RA Nguyen T.M., Wright J.R. Jr., Nielsen P.F., Conlon J.M.;
 RT "Characterization of the pancreatic hormones from the Brockmann body of the tilapia: implications for islet xenograft studies."
 RL Comp. Biochem. Physiol. 111C:33-44(1995).
 CC -1- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES THE BLOOD SUGAR LEVEL.
 CC -1- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 DR HSP; P01274; 16CN.
 DR Interpro: IPR000532; Glucagon.
 DR Pfam: PF00123; hormone2.1.
 DR PRINTS: PR00273; GLUCAGON.
 DR SMART: SM00070; GLUC; 1.
 DR PROSITE: PS00260; GLUCAGON: 1.
 KW Glucagon family: Hormone.
 SQ SEQUENCE 36 AA; 4252 MW; 5AAEF0234425756 CRC64;
 GLU1_0RENI Length: 36 December 22, 2002 19:30 Type: P Check: 382 ..
 1 HSGRTSNDY SKLEDRKAQ DRYRWLMNNK RSGABE
 11AA_SEQUENCE 1.0 STANDARD: PRT: 19 AA.
 ID LGE_ECOLI
 AC P31236;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ge1 leader peptide.
 GN GEFL OR B0018.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia
 CC Escherichia
 OK NID_TextID=562;
 RN 11AA_SEQUENCE 1.0
 RA MEDLINE=92048481; PubMed=1943701;
 RA Nielsen J.K., Kehn A., Molin S., Andersson P.;
 RT "The ge1 gene, from Escherichia coli is regulated at the level of translation."
 RT Mol. Microbiol. 5:1639-1648(1991).
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MC1655;
 RX MEDLINE=9742617; PubMed=9278503;
 RA Blatter F.R., Plunkett G., Iff Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregory J., Coats N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.
 RT "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1453-1474(1997).

CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AE000112; AAC73129.1; ALT_TERM.
 DR PIR; S16473; S16473.
 DR Leader peptide; Complete proteome.
 KW Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthomorphi; Acanthopterygii; Perciformes; Labroidae; Cichlidae; Oreochromis.
 SQ SEQUENCE 19 AA; 2259 MW; 19B3EDF371EB0BEB CRC64;
 LPE_ECOLI Length: 19 December 22, 2002 19:30 Type: P Check: 4518 ..
 1 MINCRVPLT DRVYKXRA
 11AA_SEQUENCE 1.0 STANDARD: PRT: 14 AA.
 ID MY14_ESTRO
 AC P46979;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Myosin-like tetradecapeptide (ATP).
 OS Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthomorphi; Acanthopterygii; Perciformes; Labroidae; Cichlidae; Oreochromis.
 NC NCB1_TaxID=6396;
 RX MEDLINE=96087879; PubMed=8532604;
 RA Ikeda T., Fujita T., Minakata H., Ikeda K., Omi T., Matsushima O., Ikeda T., Fujita T., Minakata H., Nomoto K.;
 RT "A novel gut tetradecapeptide isolated from the earthworm, Eisenia fetida."
 RL Peptides 16:995-999(1995).
 CC -1- FUNCTION: HAS A STIMULATIVE EFFECT ON THE CONTRACTION OF GUT MUSCLES.
 CC -1- SIMILARITY: TO INSECTS ALLATROPIN.
 KW Neuropeptide; Inhibition.
 ST MOD_RES 14 AA; 1478 MW; CC9ABEF941CD91AD CRC64;
 SQ SEQUENCE 14 AA; 1478 MW; CC9ABEF941CD91AD CRC64;
 MY14_ESTRO Length: 14 December 22, 2002 19:30 Type: P Check: 7600 ..
 1 GPKDADRI SHGF
 11AA_SEQUENCE 1.0 STANDARD: PRT: 14 AA.
 ID MY14_PHEVT
 AC P46980;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Myosin-like tetradecapeptide (ATP).
 OS Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthomorphi; Acanthopterygii; Perciformes; Labroidae; Cichlidae; Oreochromis.
 NC NCB1_TaxID=46574;
 RX MEDLINE=96087879; PubMed=8532604;
 RA Ikeda T., Fujita T., Minakata H., Ikeda K., Omi T., Matsushima O., Ikeda T., Fujita T., Minakata H., Nomoto K.;
 RT "A novel gut tetradecapeptide isolated from the earthworm, Eisenia fetida."
 RL Peptides 16:995-999(1995).
 CC -1- FUNCTION: HAS A STIMULATIVE EFFECT ON THE CONTRACTION OF GUT MUSCLES.
 CC -1- SIMILARITY: TO INSECTS ALLATROPIN.

```

KM Neuropeptide; Amidation.
FT MOD_RES 14
SQ SEQUENCE 14 AA: 1522 MW: D40DEB67CCD91AD CRC64;
MT14_PHEVI Length: 14 December 22, 2002 19:30 Type: P Check: 7729 ..

1 GFRGSDRI SHGR

!!AA_SEQUENCE 1.0 STANDARD; PRT: 23 AA.
ID T2A_PARTE
AC 027173;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trichocyst matrix protein T2-A (Secretory granule protein T2-A)
DE (TMP 2-A) (Fragment).
GN T2A.
OS Paramacium tetraurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
OC Paramacium.
OX NCBI_TaxID=5888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-D4-2;
RX MEDLINE-96059477; PubMed-7579685;
RA Madeddu L., Gautier M.-C., Vaysse L., Houari A., Sperling L.;
RT "A large multigene family codes for the polypeptides of the
RT crystalline trichocyst matrix in Paramacium."
RT Mol. Biol. Cell 6:649-659(1995).
RN [2]
RP PARTIAL SEQUENCE.
RC STRAIN-D4-2;
RX MEDLINE-95119139; PubMed-7819344;
RA Madeddu L., Gautier M.-C., le Caer J.P., de Loubresse N., Sperling L.;
RT "Protein processing and morphogenesis of secretory granules in
RT Paramacium."
RT Biochimie 76:329-335(1994).
CC -1- FUNCTION: STRUCTURAL PROTEIN THAT CRYSTALLIZE INSIDE THE
CC TRICHO CYST MATRIX.
CC -1- SUBCELLULAR LOCATION: TRICHO CYST. THESE ARE ARCHITECTURALLY
CC COMPLEX SECRETORY STORAGE GRANULES-DOCKED AT THE PLASMA MEMBRANE,
CC READY TO RAPIDLY RESPOND TO AN EXOCYTOTIC STIMULUS.
CC -1- SIMILARITY: BELONGS TO THE TMP FAMILY.
CC -1- DATABASE: NAME-Protein SpotLight;
CC NOTE-Issue 3 of October 2000;
CC WWW="http://www.expasy.org/spotlight/articles/split003.html".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U27509; AAA92609.1;
CC POLYPROTEIN; structural protein; Multigene family.
CC NON_TER 1 1
CC NON_TER 23 23
CC SEQUENCE 23 AA: 2720 MW: 08972B1F2759BDB7 CRC64;
T2A_PARTE Length: 23 December 22, 2002 19:30 Type: P Check: 946 ..

1 DPDLRLSTL TLEDRYVAE OKE

!!AA_SEQUENCE 1.0 STANDARD; PRT: 23 AA.
ID T2B_PARTE
AC 027174;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trichocyst matrix protein T2-B (Secretory granule protein T2-B)
DE (TMP 2-B) (Fragment).

```

```

GN T2B.
OS Paramacium tetraurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
OC Paramacium.
OX NCBI_TaxID=5888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-D4-2;
RX MEDLINE-96059477; PubMed-7579685;
RA Madeddu L., Gautier M.-C., Vaysse L., Houari A., Sperling L.;
RT "A large multigene family codes for the polypeptides of the
RT crystalline trichocyst matrix in Paramacium."
RT Mol. Biol. Cell 6:649-659(1995).
RN [2]
RP PARTIAL SEQUENCE.
RC STRAIN-D4-2;
RX MEDLINE-95119139; PubMed-7819344;
RA Madeddu L., Gautier M.-C., le Caer J.P., de Loubresse N., Sperling L.;
RT "Protein processing and morphogenesis of secretory granules in
RT Paramacium."
RT Biochimie 76:329-335(1994).
CC -1- FUNCTION: STRUCTURAL PROTEIN THAT CRYSTALLIZE INSIDE THE
CC TRICHO CYST MATRIX.
CC -1- SUBCELLULAR LOCATION: TRICHO CYST. THESE ARE ARCHITECTURALLY
CC COMPLEX SECRETORY STORAGE GRANULES-DOCKED AT THE PLASMA MEMBRANE,
CC READY TO RAPIDLY RESPOND TO AN EXOCYTOTIC STIMULUS.
CC -1- SIMILARITY: BELONGS TO THE TMP FAMILY.
CC -1- DATABASE: NAME-Protein SpotLight;
CC NOTE-Issue 3 of October 2000;
CC WWW="http://www.expasy.org/spotlight/articles/split003.html".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U27510; AAA92610.1;
CC POLYPROTEIN; structural protein; Multigene family.
CC NON_TER 1 1
CC NON_TER 23 23
CC SEQUENCE 23 AA: 2706 MW: 184D4B1F2759BDB7 CRC64;
T2B_PARTE Length: 23 December 22, 2002 19:30 Type: P Check: 1016 ..

1 DPDLRLSTL TLEDRYVAE OKE

!!AA_SEQUENCE 1.0 PRELIMINARY; PRT: 17 AA.
ID 09G633;
AC 09G633;
DT 01-MAY-1999 (Trembl,rel. 10, Created)
DT 01-MAY-1999 (Trembl,rel. 10, Last sequence update)
DT 01-MAY-1999 (Trembl,rel. 10, Last annotation update)
DE Subtilisin/chymotrypsin inhibitor (fragment).
DE Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-L2 434B;
RA Mang L., Steenburgh S.D., Zheng Y., Larsen S.H.;
RT "Gene identification of Chlamydia trachomatis by random DNA
RT sequencing (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB087336; AAD04110.1;
FT NON_TER 1 1
FT NON_TER 17 17
FT SEQUENCE 17 AA: 2043 MW: 4FEF704E041E120 CRC64;
09G633 Length: 17 December 22, 2002 19:30 Type: P Check: 1658 ..

```

```
1 NFGQPECTD RKNRFRI

11AA_SEQUENCE 1.0
ID O9K16 PRELIMINARY; PRT; 24 AA.
AC O9K16;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE F81r (Fragment).
GN F81r.
OS Streptomyces coelicolor A3(2).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=100226;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M145;
RA Kormanec J., Sevcikova B., Homerova D.;
RU Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF230491; AAF82064.1; -.
FT NON_TER 24
SQ SEQUENCE 24 AA; 2524 MW; F0A1B812ED67583B CRC64;

O9K16 Length: 24 December 22, 2002 19:30 Type: P Check: 2042 ..

1 MAADRGTDE ATAEALGSR APEF

11AA_SEQUENCE 1.0
ID O9X9U4 PRELIMINARY; PRT; 25 AA.
AC O9X9U4;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE ADP glucose pyrophosphorylase (Fragment).
GN GLGc.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bruton C.J.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=98062210; PubMed=9401038;
RA Martin M., Schneider D., Bruton C.J., Chater K.F., Hardisson C.;
RT A g1gc Gene Essential Only for the First of Two Spatially Distinct
RT Phases of Glycogen Synthesis in Streptomyces coelicolor A3(2).";
RL J. Bacteriol. 179:7784-7789(1997).
DR EMBL; AJ243803; CAB50742.1; -.
DR Interpro: IPR001825; NTP_transferase.
DR Pfam: PF00483; NTP_transferase; 1.
FT NON_TER 25
SQ SEQUENCE 25 AA; 2566 MW; 84B94A728A41D25C CRC64;

O9X9U4 Length: 25 December 22, 2002 19:30 Type: P Check: 4039 ..

1 MGIIVLAGE GKRLMPLTAD RAKPA

11AA_SEQUENCE 1.0
ID O49137 PRELIMINARY; PRT; 32 AA.
AC O49137;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Mxas protein (Fragment).
GN MxAS.
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
```

```
OX NCBI_TaxID=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AM1;
RX MEDLINE=96074311; PubMed=7592474;
RA Morris C.J., Kim Y.M., Perkins K.E., Lidstrom M.E.;
RT "Identification and nucleotide sequences of mxaA, mxaC, mxaX,
RT and mxaD genes from Methylobacterium extorquens AML.";
RL J. Bacteriol. 177:6825-6831(1995).
DR EMBL; L41608; AAA85567.1; -.
FT NON_TER 1
SQ SEQUENCE 32 AA; 3700 MW; EB03BB62B3C56ED5 CRC64;

O49137 Length: 32 December 22, 2002 19:30 Type: P Check: 9647 ..

1 ALRRICAPPA RPPFRLDRF DAFALSRHLM TT

11AA_SEQUENCE 1.0
ID O53914 PRELIMINARY; PRT; 9 AA.
AC O53914;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Curd protein (Fragment).
OS Streptomyces cyaneus (Streptomyces curacoi).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1904;
RN [1]
RP SEQUENCE FROM N.A.
RA Bergh S.T., Uhlen M.;
RT "Cloning, analysis and heterologous expression of the polyketides
RT synthesis genes of Streptomyces curacoi.";
RL Submitted (JUN-1990) to the EMBL/GenBank/DBJ databases.
DR EMBL; M33704; AAA26724.1; -.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1027 MW; 995BDDDC4140AB1 CRC64;

O53914 Length: 9 December 22, 2002 19:30 Type: P Check: 3396 ..

1 ITDRRAAQP

11AA_SEQUENCE 1.0
ID O50476 PRELIMINARY; PRT; 13 AA.
AC O50476;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Catalase.
GN KATG.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriinae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L10373;
RA Cockerill F.R., Uhl J.R., Temesgen Z., Zhang Y., Stockman L.,
RA Roberts G.D., Williams D.L., Kline B.C.;
RT "Rapid identification of a point mutation of the Mycobacterium
RT tuberculosis catalase-peroxidase (katG) gene associated with isoniazid
RT resistance.";
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U06263; AAB59976.1; -.
SQ SEQUENCE 13 AA; 1564 MW; 2F39A45FE994777 CRC64;

O50476 Length: 13 December 22, 2002 19:30 Type: P Check: 7152 ..

1 MPEQHPPITD RSR

11AA_SEQUENCE 1.0
ID O05422 PRELIMINARY; PRT; 38 AA.
AC O05422;
```

```

DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE Phytoene dehydrogenase (Fragment).
GN CRT1.
OS Mycobacterium marinum.
OC Bacteria: Firmicutes: Actinobacteria: Actinobacteridae:
OC Actinomycetales: Corynebacterineae, Mycobacteriaceae, Mycobacterium.
OX NCBI_TaxID=1781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M;
RX MEDLINE=97440138; PubMed=9294446;
RA Ramakrishnan L., Tran H.T., Federspiel N.A., Falkow S.;
RT "A crtb homolog essential for photochromogenicity in Mycobacterium
RT marinum: Isolation, characterization, and gene disruption via
RT homologous recombination."
RL J. Bacteriol. 179:5862-5868(1997).
DR EMBL: U92075; AAB71427.1; -.
FT NON_TER 1 1
SQ SEQUENCE 38 AA; 3986 MW; 6546332707CCDCAB CRC64;

005422 Length: 38 December 22, 2002 19:30 Type: P Check: 6757 ..

1 VPGVGVPPTL ISGRLAADRI TGNTRTSIRH LDKAKQLS

!!AA_SEQUENCE 1.0
ID 093A12 PRELIMINARY; PRT; 22 AA.
AC 093A12;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Putative ribitol/glucoase oxidoreductase (Fragment).
GN SDR2.
OS Thioacillus ferrooxidans.
OC Bacteria: Proteobacteria: gamma subdivision: Acidithiobacillus.
OX NCBI_TaxID=920;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC33020;
RA Bruscella P., Levican G., Ratouchniak J., Holmes D., Bonnefoy V.;
RT "A second operon encoding a bcl complex in Acidithiobacillus
RT ferrooxidans."
RL Submitted (Jul-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ311888; CAC44744.1; -.
FT NON_TER 1 1
SQ SEQUENCE 22 AA; 2464 MW; 3725F8E43BBA75C2 CRC64;

093A12 Length: 22 December 22, 2002 19:30 Type: P Check: 9419 ..

1 AQNFINPDL DREKDLATS TV

!!AA_SEQUENCE 1.0
ID 093A08 PRELIMINARY; PRT; 12 AA.
AC 093A08;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ResB protein (Fragment).
GN RESB.
OS Thioacillus ferrooxidans.
OC Bacteria: Proteobacteria: gamma subdivision: Acidithiobacillus.
OX NCBI_TaxID=920;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC33020;
RA Levican G., Bruscella P., Guacunanano M., Inostroza C., Jedlicki E.,
RA Bonnefoy V., Holmes D.S.;
RT "Characterization of the pet and res operons of Acidithiobacillus
RT ferrooxidans."
RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ413194; CAC88360.1; -.
FT NON_TER 1 1

```

```

SQ SEQUENCE 12 AA; 1405 MW; 886AB7DF1E13240A CRC64;

093A08 Length: 12 December 22, 2002 19:30 Type: P Check: 5753 ..

1 QSDDRKENN DG

!!AA_SEQUENCE 1.0
ID 003977 PRELIMINARY; PRT; 19 AA.
AC 003977;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE Alpha-acceptor polypeptide M15 (Fragment).
OS Escherichia coli.
OC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DH5 ALPHA;
RX MEDLINE=93083990; PubMed=1339377;
RA Prentki P.C.;
RT "Nucleotide sequence of the classical lacZ deletion delta M15."
RL Gene 122:231-232(1992).
DR EMBL: X58252; CAA41206.1; -.
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2092 MW; 8324315E003AA053 CRC64;

003977 Length: 19 December 22, 2002 19:30 Type: P Check: 4923 ..

1 MTMTIDSLAV VARTDRPSQ

!!AA_SEQUENCE 1.0
ID 099094 PRELIMINARY; PRT; 36 AA.
AC 099094;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-FEB-1997 (TREMBLrel. 02, Last annotation update)
DE MKA protein (Fragment).
OS Salmonella typhimurium.
OX Plasmid pYA426.
OC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RA Gulig P.A., Chido V.A.;
RL Infect. Immun. 58:2651-2658(1991).
DR EMBL: M64295; AAA27272.1; -.
FT NON_TER 1 1
SQ SEQUENCE 36 AA; 4121 MW; C3D43D1C622EBAB7 CRC64;

099094 Length: 36 December 22, 2002 19:30 Type: P Check: 1642 ..

1 GSDPASFQSL KLRISDDRTA DTRNRKRIIN MRVLNS

!!AA_SEQUENCE 1.0
ID 09N225 PRELIMINARY; PRT; 20 AA.
AC 09N225;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE ATP7B (Fragment).
OS Homo sapiens (Human).
OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi;
OC Mammalia: Eutheria: Primates: Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wu Z.Y., Wang N., MURONG S.X., Lin M.T., Fang L.;
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF220215; AAF67661.1; -.

```

DR InterPro: IPR001757; ATPase_E1-E2.
 DR Pfam: PF00122; E1-E2_ATPase; 1.
 FT NON_TER 1
 FT VARIANT 19 19 V -> I.
 FT NON_TER 20 20
 SO SEQUENCE 20 AA; 2282 MW; 0026BE2CE0692E3 CRC64;

Q9NZ25 Length: 20 December 22, 2002 19:30 Type: P Check: 6136 ..

1 APICOLADRF SGYFVPEIYI

!!AA_SEQUENCE 1.0
 ID 015244 PRELIMINARY; PRT; 40 AA.
 AC 015244;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE Phosphoribosylpyrophosphate synthetase isoform I (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD;
 RA Ishizuka T.;
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD;
 RA MEDLINE=92223087; PubMed=1314091;
 RA Ishizuka T., Iizasa T., Taira M., Ishijima S., Sonoda T., Shinada H.,
 RA Nagatake N., Taitbana M.;
 RT "Promoter regions of the human X-linked housekeeping genes PRPS1 and
 RT PRPS2 encoding phosphoribosylpyrophosphate synthetase subunit I and II
 RT isoforms.";
 RL Biochim. Biophys. Acta 1130:139-148(1992).
 DR EMBL: D28133; BAA05675.1; -;
 DR HSSP: P14193; IDKU.
 FT NON_TER 40 40
 SO SEQUENCE 40 AA; 4446 MW; 57C2B4011860B098 CRC64;
 Q15244 Length: 40 December 22, 2002 19:30 Type: P Check: 2868 ..

1 MPNIKIFSGS SHODLSOKIA DRLGLELKV YTKKFSNOET

!!AA_SEQUENCE 1.0
 ID 015245 PRELIMINARY; PRT; 40 AA.
 AC 015245;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE Phosphoribosylpyrophosphate synthetase isoform II (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD;
 RA Ishizuka T.;
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD;
 RA MEDLINE=92223087; PubMed=1314091;
 RA Ishizuka T., Iizasa T., Taira M., Ishijima S., Sonoda T., Shinada H.,
 RA Nagatake N., Taitbana M.;
 RT "Promoter regions of the human X-linked housekeeping genes PRPS1 and
 RT PRPS2 encoding phosphoribosylpyrophosphate synthetase subunit I and II
 RT isoforms.";
 RL Biochim. Biophys. Acta 1130:139-148(1992).
 DR EMBL: D28134; BAA05676.1; -;

FT NON_TER 40 40
 SO SEQUENCE 40 AA; 4431 MW; CEC2B31686083EEF CRC64;

Q15245 Length: 40 December 22, 2002 19:30 Type: P Check: 3314 ..

1 MPNIIVFSGS SHODLSORVA DRLGLELKV YTKKFSNOET

!!AA_SEQUENCE 1.0
 ID 09BOV8 PRELIMINARY; PRT; 34 AA.
 AC 09BOV8;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE L-isoaspartyl/D-aspartyl methyltransferase (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=94354847; PubMed=8074695;
 RA Tsai W., Clarke S.;
 RT "Amino acid polymorphisms of the human L-isoaspartyl/D-aspartyl
 RT methyltransferase involved in protein repair.";
 RL Biochem. Biophys. Res. Commun. 203:491-497(1994).
 DR EMBL: S73902; AAC60659.2; -;
 DR InterPro: IPR000682; PCMT.
 DR Pfam: PF01135; PCMT; 1.
 DR MethyItransferase; Transferase.
 FT NON_TER 1 1
 FT NON_TER 34 34
 SO SEQUENCE 34 AA; 3873 MW; EEB18E55FA9BF377 CRC64;
 Q9BOV8 Length: 34 December 22, 2002 19:30 Type: P Check: 5960 ..

1 NGIKTKDYF EVMATDRSH YAKCNFYMDS POST

!!AA_SEQUENCE 1.0
 ID 09UC11 PRELIMINARY; PRT; 15 AA.
 AC 09UC11;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Tropomyosin (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RC MEDLINE=93195352; PubMed=8450225;
 RA Das K.M., Dasgupta A., Mandal A., Geng X.;
 RT "Autoimmunity to cytoskeletal protein tropomyosin. A clue to the
 RT pathogenetic mechanism for ulcerative colitis.";
 RL J. Immunol. 150:2487-2493(1993).
 DR InterPro: IPR000533; Tropomyosin.
 DR Pfam: PF00261; Tropomyosin; 1.
 SO SEQUENCE 15 AA; 1802 MW; 7A0993CA5A54254C CRC64;

Q9UC11 Length: 15 December 22, 2002 19:30 Type: P Check: 8961 ..

1 HIAEDADRY EVVAR

!!AA_SEQUENCE 1.0
 ID 096F68 PRELIMINARY; PRT; 28 AA.
 AC 096F68;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Unknown (Protein for IMAGE:4563468) (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC011565; RAH11565.1; -
 FT NON_TER 1 1
 SQ SEQUENCE 28 AA; 2841 MW; 5799D138245D3951 CRC64;

Q96F68 Length: 28 December 22, 2002 19:30 Type: P Check: 588 ..

1 VSQPGSCRC ADRVGHVGR AGAGVRE

!!AA_SEQUENCE 1.0
 ID Q9NB1; PRELIMINARY; PRT; 26 AA.
 AC Q9NB1;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DE Stretch-MICK (Fragment).
 DR STRN-MICK OR CG8304 OR CG1825.
 GN Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Champagne M.B., Edwards K.A., Erickson H.P., Kiehart D.P.;
 RT "Drosophila stretch-MICK is a Novel Member of the Titlin/Myosin Light
 Chain Kinase Family";
 RT J. Mol. Biol. 0:0-0(2000).
 DR EMBL: AF257309; AAF50127.1; -
 DR Flybase: FBgn0013986; Strn-MICK.
 FT NON_TER 1 1
 SQ SEQUENCE 26 AA; 3004 MW; 2F2CB3A55E7FE033 CRC64;

Q9NB1 Length: 26 December 22, 2002 19:30 Type: P Check: 6662 ..

1 VVARNNGTD RLFVTITIKI PKKKE

!!AA_SEQUENCE 1.0
 ID Q9GU5; PRELIMINARY; PRT; 27 AA.
 AC Q9GU5;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE Homeobox protein Sycox2 (Fragment).
 GN SYCOX2.
 OS Sycon raphanus.
 OC Eukaryota; Metazoa; Porifera; Calcarea; Calcarenea; Leucosoleniida;
 OC Sycetiidae.
 OX NCBI_TaxID=56443;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20476455; PubMed=11020308;
 RA Manuel M., Le Parco Y.;
 RT "Homeobox Gene Diversification in the Calcareous Sponge, Sycon
 raphanus";
 RT Mol. Phylogenet. Evol. 17:97-107(2000).
 DR EMBL: AF197140; AAG28510.1; -
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox; 1.
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2986 MW; CA58FA11B629E3B7 CRC64;

Q9GU5 Length: 27 December 22, 2002 19:30 Type: P Check: 9325 ..

1 KTSKYLAPT RQALAARGL TDRQVK

!!AA_SEQUENCE 1.0

ID Q9GU4 PRELIMINARY; PRT; 27 AA.
 AC Q9GU4;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE Homeobox protein Sycox3 (Fragment).
 GN SYCOX3.
 OS Sycon raphanus.
 OC Eukaryota; Metazoa; Porifera; Calcarea; Calcarenea; Leucosoleniida;
 OC Sycetiidae.
 OX NCBI_TaxID=56443;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20476455; PubMed=11020308;
 RA Manuel M., Le Parco Y.;
 RT "Homeobox Gene Diversification in the Calcareous Sponge, Sycon
 raphanus";
 RT Mol. Phylogenet. Evol. 17:97-107(2000).
 DR EMBL: AF197141; AAG28511.1; -
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox; 1.
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 3014 MW; CA58E2803F29E3B7 CRC64;

Q9GU4 Length: 27 December 22, 2002 19:30 Type: P Check: 9332 ..

1 RTSKYLAPT RQALAARGL TDRQVK

!!AA_SEQUENCE 1.0
 ID Q9XZ6; PRELIMINARY; PRT; 23 AA.
 AC Q9XZ6;
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE Calmodulin (Fragment).
 GN CAM.
 OS Littorina saxatilis.
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Mesogastropoda; Littorinidae; Littorinidae; Littorina.
 OX NCBI_TaxID=31220;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wilding C.S., Graham J., Mill P.J.;
 RT "Molecular characterisation of calmodulin intron variation in
 Littorina (Gastropoda: Prosobranchia) species";
 RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ133335; CAB44224.1; -
 FT NON_TER 23 23
 SQ SEQUENCE 23 AA; 2815 MW; 5762B9C4E9886A06 CRC64;

Q9XZ6 Length: 23 December 22, 2002 19:30 Type: P Check: 1592 ..

1 LQDMINEVDA DRQDDRLPR VPY

!!AA_SEQUENCE 1.0
 ID Q94681; PRELIMINARY; PRT; 25 AA.
 AC Q94681;
 DT 01-FEB-1997 (TRENBLrel. 02, Created)
 DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE Pmbox1 (Fragment).
 GN Pmbox1.
 OS Polyandrocarpa misakiensis.
 OC Eukaryota; Metazoa; Chordata; Trochordata; Ascidiacea;
 OC Stolidobranchia; Styelidae; Polyandrocarpa.
 OX NCBI_TaxID=7723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USHIMADO.
 RA Fujiwara S., Kawamura K.;
 RT "Cloning of homeobox-containing genes from the budding ascidian
 Polyandrocarpa misakiensis";


```

RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: D88274; BAA13570.1; -.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
FT NON_TER 1
SQ SEQUENCE 25 AA; 3074 MW; D7CBE263FB30A51 CRC64;
Q94681 Length: 25 December 22, 2002 19:30 Type: P Check: 5282 ..

1 HFNQYLCRR ROYAKAVNL TDROY
11AA_SEQUENCE 1.0
ID Q23749 PRELIMINARY; PRT; 27 AA.
AC Q23749;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Oryz ortholog homeobox (Fragment).
GN CTS-OX1.
OS Ctenodrilus serratus.
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;
OC Terebellida; Ctenodrilidae; Ctenodrilus.
OX NCBI_TaxID=40316;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94356262; PubMed=7915607;
RA Dick M.H., Buss L.W.;
RT "A PCR-based survey of homeobox genes in Ctenodrilus serratus
RT (Annelida: Polychaeta).";
RL Mol. Phylogenet. Evol. 3:146-158(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Dick M.H., Buss L.W.;
RT Submitted (MAY-1995) to the EMBL/Genbank/DBJ databases.
RL EMBL: U26634; AAC46856.1; -.
RW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
SQ SEQUENCE 27 AA; 2879 MW; 4CCE22303AA0102 CRC64;
Q23749 Length: 27 December 22, 2002 19:30 Type: P Check: 8475 ..

1 PQEIVTDC TIAHRASPET DRSGEN
11AA_SEQUENCE 1.0
ID Q25482 PRELIMINARY; PRT; 27 AA.
AC Q25482;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Anthox1.MS (Fragment).
GN ANTHOX1.MS.
OS Metridium senile (Brown sea anemone) (Frilled sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actinaria;
OC Myantheae; Metridiidae; Metridium.
OX NCBI_TaxID=6116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97435315; PubMed=9290214;
RA Finerly J.R., Martindale M.Q.;
RT "Homeoboxes in sea anemones (Cnidaria:Anthozoa): a PCR-based survey of
RT Nematostella vectensis and Metridium senile.";
RL Biol. Bull. 193:62-76(1997).
DR EMBL: U42727; AAA86626.1; -.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00024; HOMEBOX.
FT NON_TER 1
SQ SEQUENCE 27 AA; 3314 MW; E6B0D191EE3905CF CRC64;
Q25482 Length: 27 December 22, 2002 19:30 Type: P Check: 9269 ..

```

```

1 HFNHETKR RSEMATQNL TDROYKI
11AA_SEQUENCE 1.0
ID Q94458 PRELIMINARY; PRT; 27 AA.
AC Q94458;
DT 01-FEB-1997 (TREMblrel. 02, Created)
DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Chv-Hb11 protein (Fragment).
GN Chv-Hb11.
OS Chaetopterus variopedatus.
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;
OC Spionida; Chaetopteridae; Chaetopterus.
OX NCBI_TaxID=34590;
RN [1]
RP SEQUENCE FROM N.A.
RA Irvine S.M., Mariner S.A., Hunter J.D., Martindale M.Q.;
RT "A survey of homeobox genes in Chaetopterus variopedatus and analysis
RT of polychaete homeodomain.";
RL Submitted (AUG-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: U68283; AAB16992.1; -.
FT NON_TER 1
SQ SEQUENCE 27 AA; 3351 MW; 74E5C77ED08CDEE4 CRC64;
Q94458 Length: 27 December 22, 2002 19:30 Type: P Check: 8578 ..

1 NOKRFEKKD RDRISNEIGL DDRQIKY
11AA_SEQUENCE 1.0
ID Q96892 PRELIMINARY; PRT; 27 AA.
AC Q96892;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Caudal (Fragment).
GN CAD.
OS Sacculina carcini.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Maxillopoda; Cirripedia;
OC Rhizocephala; Kentrogonida; Sacculinidae; Sacculina.
OX NCBI_TaxID=51650;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98334561; PubMed=9667986;
RA Mouchel-Vielh E., Rigolot C., Gilbert J.M., Deutsch J.S.;
RT "Molecules and the body plan: the Hox genes of Cirripedes
RT (Crustacea).";
RL Mol. Phylogenet. Evol. 9:382-389(1998).
DR EMBL: U79471; AAD00342.1; -.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00024; HOMEBOX.
FT NON_TER 1
SQ SEQUENCE 27 AA; 3289 MW; 1CDCE2E4CCA04E00F CRC64;
Q96892 Length: 27 December 22, 2002 19:30 Type: P Check: 9212 ..

1 RFNNTTKR KLEISRTGL TDROYKI
11AA_SEQUENCE 1.0
ID P91901 PRELIMINARY; PRT; 27 AA.
AC P91901;
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Homeobox protein Bhox26 (Fragment).
OS Berce ovata.
OC Eukaryota; Metazoa; Ctenophora; Cycliocoela; Beroidea; Beroidea; Berce.
OX NCBI_TaxID=10201;
RN [1]
RP SEQUENCE FROM N.A.

```

RA Mitchel H.J., Meech R.W.;
 RT "Bbox26 - Berce Homeobox";
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U89381; AAB49471.1; -
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox; 1.
 DR PRINTS: PR00024; HOMEBOX.
 DR ProDom: PD000010; Homeobox; 1.
 FT NON_TER 1 1
 FT SEQUENCE 27 27
 SO SEQUENCE 27 AA; 338 MW; 55318651A81607BD CRC64;
 P91901 Length: 27 December 22, 2002 19:30 Type: P Check: 9410 ..

1 LFNNYLTRER RLEISKINL TDRQVKI
 11AA_SEQUENCE 1.0
 ID 002602; PRELIMINARY; PRT; 27 AA.
 AC 002602; 002535; -
 DT 01-JUL-1997 (TEMBLrel. 04, Created)
 DT 01-JUL-1997 (TEMBLrel. 04, Last sequence update)
 DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
 DE Homeobox protein Bbox35 (Fragment).
 OS Berce ovata.
 OC Eukaryota; Metazoa; Ctenophora; Cyclocoela; Beroidea; Beroidea; Berce.
 OX NCBI_TaxID=10201;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mitchel H.J., Meech R.W.;
 RT "Bbox35 - Berce Homeobox";
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U89383; AAB49473.1; -
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox; 1.
 DR PRINTS: PR00024; HOMEBOX.
 DR ProDom: PD000010; Homeobox; 1.
 FT NON_TER 1 1
 FT SEQUENCE 27 27
 SO SEQUENCE 27 AA; 332 MW; 5529340A951607BD CRC64;
 002602 Length: 27 December 22, 2002 19:30 Type: P Check: 9552 ..

1 LFNNYLTRER RLEISGVNL TDRQVKI
 11AA_SEQUENCE 1.0
 ID 018615; PRELIMINARY; PRT; 34 AA.
 AC 018615;
 DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE Ftz-rl (Fragment).
 OS Artemia salina (Brine shrimp).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
 OC Artemiidae; Artemia.
 OX NCBI_TaxID=85549;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=97338099; PubMed=9192646;
 RA Escrivá H., Safi R., Hanni C., Langlois M.C., Saumitou-Laprade P.,
 RA Schenelin D., Capron A., Pierce R., Laudet V.;
 RT "Ligand binding was acquired during evolution of nuclear receptors";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:6803-6808(1997).
 CC - SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC - SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 DR EMBL: U93418; AAC83395.1; -
 DR InterPro: IPR001628; Znf_C4steroid.
 DR Pfam: PF00105; Zf-C4; 1.
 DR ProDom: PD000035; Znf_C4steroid; 1.
 DR SMART: SM00399; Znf_C4; 1.
 KM DNA-binding; Nuclear protein; Receptor; Transcription regulation;
 FT zinc-finger.
 FT NON_TER 1 1
 FT SEQUENCE 34 34

1 FKRTRVONKRV YTCMADRSCN IDKSQRRRCPC PCRF
 11AA_SEQUENCE 1.0
 ID 09Y973; PRELIMINARY; PRT; 17 AA.
 AC 09Y973;
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
 DE CG12485 protein.
 GN CG12485.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaratunga C., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V., Chao M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abil J.F., Adyanthi A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Boltskov S.,
 RA Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke G., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.U., Mei M.-H., Idegawa C.,
 RA Jaisl M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Jia L.M., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003791; AAF57423.1; -
 DR FlyBase: FBgn0034547; CG12485.
 SO SEQUENCE 17 AA; 1907 MW; 47D598D29F4860DE CRC64;
 Q9Y973 Length: 17 December 22, 2002 19:30 Type: P Check: 1645 ..

1 MELSTDRGP HHAQOGR
 11AA_SEQUENCE 1.0
 ID 09Y971; PRELIMINARY; PRT; 21 AA.
 AC 09Y971;
 DT 01-MAY-2000 (TEMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, last sequence update)
 DT 01-JUN-2000 (TReMBLrel. 14, last annotation update)
 DE Cysteine proteinase (Fragment).
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Bombycoidea; Bombycidae; Bombyx.
 OX NCBI_TaxID=7091;
 RN [1]
 RP MEDLINE=94086476; PubMed=8262908;
 RA Takahashi S.Y., Yamamoto Y., Shiono Y., Kageyama T.;
 RT Cysteine proteinase from the eggs of the silkworm, Bombyx mori;
 RT Identification of a latent enzyme and characterization of activation
 RT and proteolytic processing in vivo and in vitro.";
 RL J. Biochem. 114:267-272(1993).
 SQ SEQUENCE 21 AA; 2379 MW; 9232784B6DAEAB1 CRC64;

09TW1 Length: 21 December 22, 2002 19:30 Type: P Check: 8106 ..

1 XPROVDPRKH GAVTDKXXQ X

!!AA_SEQUENCE 1.0
 ID 062575 PRELIMINARY; PRT; 26 AA.
 AC 062575;
 DT 01-AUG-1998 (TReMBLrel. 07, Created)
 DT 01-OCT-2000 (TReMBLrel. 07, last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, last annotation update)
 DE Hox protein MYX4 (Fragment).
 GN HOX OR MYX4.
 OS Tetracapsula hyzoides.
 OC Eukaryota; Metazoa; Myxozoa; Myxozoa incertae sedis; Tetracapsula.
 OX NCBI_TaxID=75694;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98196665; PubMed=9537319;
 RA Anderson C.L., Canning E.U., Okamura B.;
 RT "A triploblast origin for Myxozoa?";
 RL Nature 392:346-347(1998).
 DR EMBL: AJ005124; CA06387.1; -;
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox; 1.
 DR ProDom: PD000010; Homeobox; 1.
 DR DNA-binding; Homeobox; Nuclear protein.
 KW NON_TER 1
 FT NON_TER 26
 FT SEQUENCE 26 AA; 3212 MW; E38651AD1657BD2C CRC64;

062575 Length: 26 December 22, 2002 19:30 Type: P Check: 7242 ..

1 LFNWLSRER RLEISKSIDL TDROYK

!!AA_SEQUENCE 1.0
 ID 061676 PRELIMINARY; PRT; 27 AA.
 AC 061676;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, last annotation update)
 DE HOX11-D125 protein (D125) (Fragment).
 GN HOX11-D125 OR D125.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93281593; PubMed=8099440;
 RA Dear T.N., Sanchez-Garcia I., Rabbits T.H.;
 RT The HOX11 gene encodes a DNA-binding nuclear transcription factor
 RT belonging to a distinct family of homeobox genes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:4431-4435(1993).
 DR EMBL: L08618; AAA28612.1; -;

DR FlyBase: FBgn0010394; Hox11-D125.
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox; 1.
 DR ProDom: PD000010; Homeobox; 1.
 FT NON_TER 1
 FT NON_TER 27
 FT NON_TER 27
 SQ SEQUENCE 27 AA; 2937 MW; CABBC2F876F2B3A3 CRC64;

061676 Length: 27 December 22, 2002 19:30 Type: P Check: 8698 ..

1 LYQKYSPPAD RDEIATSLGL SNAQVIT

!!AA_SEQUENCE 1.0
 ID 09MZ8 PRELIMINARY; PRT; 29 AA.
 AC 09MZ8;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, last annotation update)
 DE Von Willebrand factor (Fragment).
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20072937; PubMed=10603266;
 RA Chaves R., Sampaio I., Schneider M.P., Schneider H., Page S.L.,
 RA Goodman M.;
 RT "The place of Callimico goeldii in the callitrichine phylogenetic
 RT tree: evidence from von Willebrand factor gene intron II sequences.";
 RL Mol. Phylogenet. Evol. 13:392-404(1999).
 DR EMBL: AF092833; AAF7601.1; -;
 FT NON_TER 1
 FT NON_TER 29
 FT NON_TER 29
 SQ SEQUENCE 29 AA; 3227 MW; 60F1190C8F227D03 CRC64;

09MZ8 Length: 29 December 22, 2002 19:30 Type: P Check: 2931 ..

1 GRDCQDHSFS IVIEYQCAD DRDAVCTRS

!!AA_SEQUENCE 1.0
 ID 029394 PRELIMINARY; PRT; 23 AA.
 AC 029394;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)
 DE Growth hormone receptor (Fragment).
 GN GHR.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97049323; PubMed=8894053;
 RA Venia P.J., Brouillette J.A., Yuzbasizyan-Gurkan V., Brewer G.J.;
 RT "Gene-specific universal mammalian sequence-tagged sites: application
 RT to the canine genome.";
 RL Biochem. Genet. 34:321-341(1996).
 DR EMBL: L77673; AAA97423.1; -;
 KW Receptor.
 FT NON_TER 1
 FT NON_TER 23
 FT NON_TER 23
 SQ SEQUENCE 23 AA; 2620 MW; 85442F0EC46D5D08 CRC64;

029394 Length: 23 December 22, 2002 19:30 Type: P Check: 807 ..

1 DDUDEKTBSS DTDRLSNDH EKS

!!AA_SEQUENCE 1.0
 ID 09TRS7 PRELIMINARY; PRT; 24 AA.
 AC 09TRS7;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Vitronectin (Fragment).
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN (1)
 RP SEQUENCE.
 RX MEDLINE=92207982; PubMed=1372829;
 RA Nakashima N., Miyazaki K., Ishikawa M., Yatcho T., Ogawa H.,
 Uchibori H., Matsumoto I., Seno N., Hayashi M.;
 RT "Vitronectin diversity in evolution but uniformity in ligand binding
 and size of the core polypeptide."
 RL Biochim. Biophys. Acta 1120:1-10(1992).
 DE InterPro: IPR000886; ER-target.
 DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
 FT NON_TER 1 1
 FT NON_TER 24 24
 SQ SEQUENCE 24 AA; 2745 MW; 94F0054986FE1215 CRC64;

Q9TR57 Length: 24 December 22, 2002 19:30 Type: P Check: 2610 ..

1 A0ESXKGRVT EGFNDRKQO QDEL

ID 09T009 PRELIMINARY; PRT; 15 AA.
 AC 09T009;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE Glutamate dehydrogenase isoform I (EC 1.4.1.2) (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN (1)
 RP SEQUENCE.
 RX MEDLINE=96061967; PubMed=7588764;
 RA Cho S.W., Lee J., Choi S.Y.;
 RT "Two soluble forms of glutamate dehydrogenase isoproteins from bovine
 brain."
 RL Eur. J. Biochem. 233:340-346(1995).
 RN (2)
 RP SEQUENCE.
 RX MEDLINE=96043916; PubMed=7581004;
 RA Lee J., Kim S.W., Cho S.W.;
 RT "A novel glutamate dehydrogenase from bovine brain: purification and
 characterization."
 RL Biochem. Mol. Biol. Int. 36:1087-1096(1995).
 SQ SEQUENCE 15 AA; 1754 MW; 65F7CD91023AEBA CRC64;

Q9T0Q9 Length: 15 December 22, 2002 19:30 Type: P Check: 8644 ..

1 EEAADREDD PNFEK

ID 09TR40 PRELIMINARY; PRT; 15 AA.
 AC 09TR40;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE Glutamate dehydrogenase isoform II (EC 1.4.1.2) (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN (1)
 RP SEQUENCE.
 RX MEDLINE=96061967; PubMed=7588764;
 RA Cho S.W., Lee J., Choi S.Y.;

RT "Two soluble forms of glutamate dehydrogenase isoproteins from bovine
 brain."
 RL Eur. J. Biochem. 233:340-346(1995).
 SQ SEQUENCE 15 AA; 1724 MW; 65F7CD91023AE925 CRC64;

Q9TR40 Length: 15 December 22, 2002 19:30 Type: P Check: 8661 ..

1 VEAADREDD PNFEK

ID 09TRF2 PRELIMINARY; PRT; 33 AA.
 AC 09TRF2;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 21, Last annotation update)
 DE VAMP/synaptobrevin-2 (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN (1)
 RP SEQUENCE.
 RX MEDLINE=93374072; PubMed=8365494;
 RA Horikawa H.P., Saisu H., Ishizuka T., Sekine Y., Tsugita A., Odani S.,
 Abe T.;
 RL FEBS Lett. 330:236-240(1993).
 DR InterPro: IPR001388; Synaptobrevin.
 DR Pfam: PF00957; Synaptobrevin.1.
 DR Prodom: PD001229; Synaptobrevin.1.
 DR PROSITE: PS00417; SYNAPTOBREVIN.1.
 SQ SEQUENCE 33 AA; 3672 MW; 58438688353EC55 CRC64;

Q9TR2 Length: 33 December 22, 2002 19:30 Type: P Check: 1454 ..

1 XYNVDKVLER DQKLELDR ADALAGASQ FET

ID 037112 PRELIMINARY; PRT; 22 AA.
 AC 037112;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBlrel. 08, Last annotation update)
 DE ORF22.
 OS Pinus thunbergii (Green pine) (Japanese black pine).
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 OX NCBI_TaxID=3350;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92212283; PubMed=1557027;
 RA Tsudzuki J., Nakashima K., Tsudzuki T., Hiratsuka J., Shibata M.,
 Wakasugi T., Sugitara M.;
 RT "Chloroplast DNA of black pine retains a residual inverted repeat
 lacking rRNA genes: nucleotide sequences of trnQ, trnK, psbA, trnI and
 trnH and the absence of rps16."
 RL Mol. Gen. Genet. 232:206-214(1992).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95094312; PubMed=8001170;
 RA Tsudzuki J., Ito S., Tsudzuki T., Wakasugi T., Sugitara M.;
 RT "A new gene encoding tRNA(Pro) (GGG) is present in the chloroplast
 genome of black pine: a compilation of 32 tRNA genes from black pine
 chloroplasts."
 RL Curr. Genet. 26:153-158(1994).
 RN (3)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95024047; PubMed=7937893;
 RA Wakasugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T.,
 Sugitara M.;
 RT "Loss of all ndh genes as determined by sequencing the entire
 chloroplast genome of the black pine Pinus thunbergii."

```

RL Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).
RN {}
RP SEQUENCE FROM N.A.
RA Sugliura M.;
RL Submitted (SEP-1994) to the EMBL/Genbank/DBJ databases.
RN {}
RP SEQUENCE FROM N.A.
RA MEDLINE=94138245; PubMed=8305874;
RL J.J. Goldschmidt-Clermont M., Timko M.;
RT "Chloroplast-encoded chlB is required for light-independent
  protochlorophyllide reductase activity in Chlamydomonas reinhardtii.";
RL Plant Cell 5:1817-1829(1993).
DR EMBL; D17510; BAA04307.1; -.
DR EMBL; D11467; BAA02023.1; -.
KW Chloroplast.
SQ SEQUENCE 22 AA; 2606 MW; 0A95CB0443BCAEC5 CRC64;

Q97112 Length: 22 December 22, 2002 19:30 Type: P Check: 9119 ..

1 MEXYITDRSI ECGIYIKKIE SI
11AA_SEQUENCE 1.0
ID Q95774 PRELIMINARY; PRT; 22 AA.
AC Q95774;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE Ribosomal protein S12 (Fragment).
GN RPS12.
OS Abies alba (Edelstanne) (European silver fir).
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
CX NCBI_TaxID=45372;
RN {}
RP SEQUENCE FROM N.A.
RA Liepelt S.;
RT "Sequence analysis of chosen regions of the mitochondrial genome of
  Abies alba.";
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF359456; AAK48935.1; -.
KW Mitochondrion.
FT NON_TER 22
SQ SEQUENCE 22 AA; 2711 MW; 97BF270D8C4E6558 CRC64;

Q95774 Length: 22 December 22, 2002 19:30 Type: P Check: 9875 ..

1 MPTSNQSI RH GREKKRRTRD TR
11AA_SEQUENCE 1.0
ID Q95770 PRELIMINARY; PRT; 23 AA.
AC Q95770;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Ribosomal protein S12 (Fragment).
GN RPS12.
OS Pinus mugo.
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
CX NCBI_TaxID=28528;
RN {}
RP SEQUENCE FROM N.A.
RA Machowiak W., Liepelt S., Prus-Glowacki W.;
RT "Sequence analysis of chosen regions of the mitochondrial genome of
  Pinus mugo and P. sylvestris.";
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF359576; AAK39117.1; -.
KW Mitochondrion.
FT NON_TER 23
SQ SEQUENCE 23 AA; 2782 MW; 17A7BF270D8C4E65 CRC64;

```

Q95770 Length: 23 December 22, 2002 19:30 Type: P Check: 1370 ..

1 MPTSNQSI RH GREKKRRTRD TRA

```

11AA_SEQUENCE 1.0
ID Q95758 PRELIMINARY; PRT; 23 AA.
AC Q95758;
DT 01-DEC-2001 (TREMBlrel. 13, Created)
DT 01-DEC-2001 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 16, Last annotation update)
DE Ribosomal protein S12 (Fragment).
GN RPS12.
OS Pinus sylvestris (Scots pine).
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
CX NCBI_TaxID=3349;
RN {}
RP SEQUENCE FROM N.A.
RA Machowiak W., Liepelt S., Prus-Glowacki W.;
RT "Sequence analysis of chosen regions of the mitochondrial genome of
  Pinus mugo and P. sylvestris.";
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF359577; AAK39119.1; -.
KW Mitochondrion.
FT NON_TER 23
SQ SEQUENCE 23 AA; 2782 MW; 17A7BF270D8C4E65 CRC64;

Q95758 Length: 23 December 22, 2002 19:30 Type: P Check: 1370 ..

1 MPTSNQSI RH GREKKRRTRD TRA
11AA_SEQUENCE 1.0
ID Q972K6 PRELIMINARY; PRT; 20 AA.
AC Q972K6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE Acyl-(acyl-carrier protein) hydrolase 33 kDa polypeptide, AHL
  (EC 3.1.2.14) (Fragment).
GN Cucurbita moschata (Cushaw squash) (Winter crookneck squash).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
  OC eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
CX NCBI_TaxID=3662;
RN {}
RP SEQUENCE.
RX MEDLINE=93004473; PubMed=1391766;
RA Imai H., Nishida I., Morata N.;
RT "Acyl-(acyl-carrier protein) hydrolase from squash cotyledons specific
  to long-chain fatty acids: purification and characterization.";
RL Plant Mol. Biol. 20:199-206(1992).
SQ SEQUENCE 20 AA; 1973 MW; AE5304F9CD73D36E CRC64;

Q972K6 Length: 20 December 22, 2002 19:30 Type: P Check: 5922 ..

1 GSSSLADRLX LGSLADGFS
11AA_SEQUENCE 1.0
ID Q972H3 PRELIMINARY; PRT; 24 AA.
AC Q972H3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Chaperonin 21 (Fragment).
GN Spinacia oleracea (Spinach).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
CX NCBI_TaxID=3562;
RN {}

```

RP SEQUENCE
 RX MEDLINE=95394976; PubMed=7665625;
 RA Ryan M.T., Naylor D.J., Hoogenraad N.J., Hoj P.B.;
 RL J. Biol. Chem. 270:22037-22043(1995).
 SQ SEQUENCE 24 AA; 2642 MW; B797841E1005A51A CRC64;
 Q9T2H3 Length: 24 December 22, 2002 19:30 Type: P Check: 2853 ..

1 ATVVAPKYS IKPTADRLVLI KIKE
 IIAA_SEQUENCE 1.0
 ID Q37852 PRELIMINARY; PRT; 18 AA.
 AC Q37852;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DE A protein (Fragment).
 OS Bacteriophage R17.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Leviviridae;
 OC Levivirus.
 OX NCBI_TaxID=12026;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=74125852; PubMed=4361645;
 RA Rensing U.F.E., Coulson A., Schoenmakers J.G.G.;
 RT "A sequence of 54 nucleotides from the A-protein cistron of Coliphage-
 RT R17 RNA."
 RL Eur. J. Biochem. 41:431-438(1974).
 DR EMBL; M24814; AAA32177.1; .
 FT NON_TER 1 1
 FT NON_TER 18 18
 SQ SEQUENCE 18 AA; 2217 MW; CE3DE9FD852F11E CRC64;
 Q37852 Length: 18 December 22, 2002 19:30 Type: P Check: 3123 ..

1 ALRTALNED KKFRRSHV
 IIAA_SEQUENCE 1.0
 ID Q42209 PRELIMINARY; PRT; 28 AA.
 AC Q42209;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DE Ubiquitin conjugating enzyme UBC10 (Fragment).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CV. COLUMBIA; TISSUE=SEEDLING;
 RA Desprez T., Amselem J., Chapello H., Caboche M., Hofte H.;
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z29145; CAA82396.1; .
 DR HSSP; P15731; IQCO.
 DR InterPro; IPR000608; UBO_conjugat.
 DR Pfam; PF00179; UO_con.1.
 DR Prodom; PD000461; UBO_conjugat.1.
 FT NON_TER 1 1
 FT NON_TER 28 28
 SQ SEQUENCE 28 AA; 3321 MW; C4000F6BD329NAF CRC64;
 Q42209 Length: 28 December 22, 2002 19:30 Type: P Check: 1580 ..

1 PEIAHMYKTD RAKYESTARS WTOKYAMG
 IIAA_SEQUENCE 1.0
 ID Q9FZP4 PRELIMINARY; PRT; 29 AA.
 AC Q9FZP4;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE Genomic DNA, chromosome 5, p1 clone:WNP2.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=COLUMBIA;
 RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI."
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB025636; BAB11485.1; .
 SQ SEQUENCE 29 AA; 3557 MW; 9B75DC6B293119A2 CRC64;
 Q9FZP4 Length: 29 December 22, 2002 19:30 Type: P Check: 3486 ..

1 MDVDVDFST LISKEDREN EDREPKYVY
 IIAA_SEQUENCE 1.0
 ID Q39633 PRELIMINARY; PRT; 23 AA.
 AC Q39633;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DE Catalase (Fragment).
 OS Cucumis sativus (Cucumber).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucotsids I; Cucurbitales; Cucurbitaceae; Cucumis.
 OX NCBI_TaxID=3659;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=COTYLEDON;
 RX MEDLINE=96104306; PubMed=8564304;
 RA Toyama T., Teramoto H., Takeba G., Tsuji H.;
 RT "Cytokinin induces a rapid decrease in the levels of mRNAs for
 RT catalase, 3-hydroxy-3-methylglutaryl CoA reductase, leucine and other
 RT unidentified proteins in etiolated cotyledons of cucumber.";
 RL Plant Cell Physiol. 36:1349-1359(1995).
 DR EMBL; D63385; BAA09701.1; .
 FT NON_TER 1 1
 FT NON_TER 23 23
 SQ SEQUENCE 23 AA; 2445 MW; DD39DAFD56C3AE5 CRC64;
 Q39633 Length: 23 December 22, 2002 19:30 Type: P Check: 1444 ..

1 NGSQADRSVG OKLAHNLWR PSI
 IIAA_SEQUENCE 1.0
 ID P82195 PRELIMINARY; PRT; 26 AA.
 AC P82195;
 DT 01-JUN-2000 (Tremblrel. 14, Created)
 DT 01-JUN-2000 (Tremblrel. 14, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE Chloroplast 50S ribosomal protein L18 beta (Fragment).
 OS Spinacia oleracea (Spinach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE.
 RX STRAIN=CV. ALVARO; TISSUE=LEAF;
 RX MEDLINE=20435798; PubMed=10874046;
 RA Yamauchi K., Subramanian A.R.;
 RT "The plastid ribosomal proteins. Identification of all the proteins in
 RT the 50 S subunit of an organelle ribosome (chloroplast).";
 RL J. Biol. Chem. 275:28466-28482(2000).
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 5S RIBOSOMAL RNA.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
 CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 11.5 KDA.
 CC -1- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.

KW Ribosomal protein: Chloroplast: rRNA-binding.
 FT NON_TER 26
 SO SEQUENCE 26 AA: 3184 MW: C5664ECLB2NB7C97 CRC64:

P82195 Length: 26 December 22, 2002 19:30 Type: P Check: 7110 ..

1 KAHRREDRT ARHVRIRKKV EGTPEER

!!AA_SEQUENCE 1.0
 ID P82196 PRELIMINARY: PRT: 29 AA.

AC P82196;
 DT 01-JUN-2000 (TREMblrel. 14, Created)
 DT 01-JUN-2000 (TREMblrel. 14, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE Chloroplast 50S ribosomal protein L18 alpha (Fragment).
 OS Spinacia oleracea (Spinach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Caryophyllales; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
 NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE.
 RC STRAIN-CV. ALMARO; TISSUE-LEAF;
 RX MEDLINE=20435798; PubMed=10874046;
 RA Yamauchi K., Subramanian A.R.;
 RT "The plastid ribosomal proteins. Identification of all the proteins in
 the 50 S subunit of an organelle ribosome (chloroplast).";
 RL J. Biol. Chem. 275:28466-28482(2000).
 CC - FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 5S RIBOSOMAL RNA.
 CC - SUBCELLULAR LOCATION: CHLOROPLAST.
 CC - TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
 CC - MISCELLANEOUS: ON THE 2D-GEL, ITS MW IS: 11.5 KDA.
 CC - SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.
 KM Ribosomal protein; Chloroplast; rRNA-binding.
 FT NON_TER 29
 SO SEQUENCE 29 AA: 3517 MW: 86207D55664EC1B2 CRC64:

P82196 Length: 29 December 22, 2002 19:30 Type: P Check: 4154 ..

1 KAHRREDRT ARHVRIRKKV EGTPEERXLL

!!AA_SEQUENCE 1.0
 ID P83089 PRELIMINARY: PRT: 19 AA.

AC P83089;
 DT 01-OCT-2001 (TREMblrel. 18, Created)
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Thylakoid lumenal 25.3 kDa protein (Fragment).
 OS Spinacia oleracea (Spinach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
 NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE, SUBCELLULAR LOCATION, AND MASS SPECTROMETRY.
 RC TISSUE-LEAF;
 RX PubMed=11719511;
 RA Schubert M., Petersson U.A., Haas B.J., Funk C., Schroeder W.P.,
 Kieselbach T.;
 RT "Proteome map of the chloroplast lumen of Arabidopsis thaliana";
 RL J. Biol. Chem. 277:8354-8365(2002).
 CC - SUBCELLULAR LOCATION: CHLOROPLAST; WITHIN THE THYLAKOID LUMEN.
 CC - MASS SPECTROMETRY: MW=25300; METHOD=MALDI.
 KW Chloroplast; Thylakoid.
 FT NON_TER 19
 SO SEQUENCE 19 AA: 2093 MW: EC3J42F8633F1F9 CRC64:

P83089 Length: 19 December 22, 2002 19:30 Type: P Check: 4307 ..

1 A1ANAPLLDT T1TDVFEFD

!!AA_SEQUENCE 1.0
 ID Q9S8D2 PRELIMINARY: PRT: 29 AA.

AC Q9S8D2;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE CMT1-B-TRYPSIN inhibitor.
 OS Cucumis melo (Muskmelon).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 NCBI_TaxID=3656;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=96015144; PubMed=8537309;
 RA Lee C.F., Lin J.Y.;
 RT "Amino acid sequences of trypsin inhibitors from the melon Cucumis
 melo";
 RL J. Biochem. 118:18-22(1995).
 DR HSSP: P12071; 2ETL.
 DR InterPro: IPR000737; Squash.
 DR Pfam: PF00239; Squash.1.
 DR PRINTS: PR00293; SQUASHINHEFR.
 DR PRODOM: PD003401; Squash.1.
 DR SMART: SM00286; PRT: 1.
 DR PROSITE: PS00286; SQUASH. INHIBITOR. 1.
 SO SEQUENCE 29 AA: 3196 MW: A4BCFE7AA1AC300E CRC64:

Q9S8D2 Length: 29 December 22, 2002 19:30 Type: P Check: 2605 ..

1 VGCPRILMKC KTDRCCLTNC TCKRNGYCG

!!AA_SEQUENCE 1.0
 ID Q9S898 PRELIMINARY: PRT: 16 AA.
 AC Q9S898;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE 23 kDa heat-induced protein (Fragment).
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96351184; PubMed=8742333;
 RA Sabehat A., Weiss D., Lurie S.;
 RT "The correlation between heat-shock protein accumulation and
 RT persistence and chilling tolerance in tomato fruit.";
 RL Plant Physiol. 110:531-537(1996).
 SO SEQUENCE 16 AA: 1840 MW: 980D3F77BFEB6CE3 CRC64:

Q9S898 Length: 16 December 22, 2002 19:30 Type: P Check: 484 ..

1 NNTQMTAYD QDDRGX

!!AA_SEQUENCE 1.0
 ID Q9QVE2 PRELIMINARY: PRT: 12 AA.
 AC Q9QVE2;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE TRANSFERIN-PEPTIDE 21 (Fragment).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10118;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=92165927; PubMed=1791188;
 RA Cavanaugh P.G., Nicolson G.L.;
 RT "Lung-derived growth factor that stimulates the growth of lung-
 RT metastasizing tumor cells: identification as transferrin.";
 RL J. Cell. Biochem. 47:261-271(1991).

FT NON_TER 1 1
 FT NON_TER 12 12
 SO SEQUENCE 12 AA: 1436 MW: 40AD1DFA420AADD3 CRC64;

09QVE2 Length: 12 December 22, 2002 19:30 Type: P Check: 5863 ..

1 NLANKADRDQ YE

11AA_SEQUENCE 1.0
 ID Q9CS32 PRELIMINARY: PRT: 18 AA.

AC Q9CS32;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE B30218P12R1K protein (Fragment).

OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=CORPORA OVARIUM; GEMINA;
 RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa K., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita W., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohsaki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection."
 RT Nature 409:685-690(2001).
 RL EMBL: AK019213; BAB31604.1;
 DR MGD: MGI:1914299; 2610020J05R1K.
 FT NON_TER 18 18
 SO SEQUENCE 18 AA: 2136 MW: 72A0BE772BCA48C8 CRC64;

09CTNO Length: 18 December 22, 2002 19:30 Type: P Check: 2989 ..

1 MPYKNDTR ALSLEBY

11AA_SEQUENCE 1.0
 ID Q9CS32 PRELIMINARY: PRT: 27 AA.

AC Q9CS32;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE 2610020J05R1K protein (Fragment).

OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
 RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa K., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita W., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohsaki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection."
 RT Nature 409:685-690(2001).
 RL EMBL: AK019213; BAB31604.1;
 DR MGD: MGI:1914299; 2610020J05R1K.
 FT NON_TER 27 27
 SO SEQUENCE 27 AA: 3299 MW: 95C1A07254A2C489 CRC64;

09CS32 Length: 27 December 22, 2002 19:30 Type: P Check: 8838 ..

1 MAENTDRNQ EKLNRVREL EOEVERL

11AA_SEQUENCE 1.0
 ID Q63990 PRELIMINARY: PRT: 25 AA.

AC Q63990;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Hox3.5 homeobox homolog protein (Fragment).

OS Rattus norvegicus (Rat)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=MISTAR;
 TX TISSUE=BONE MORPHOGENETIC PROTEIN-IMPLANTED SUBCUTANEOUS MUSCLE;
 RX MEDLINE=9427162; PubMed=7911662;

RA Timura T., Oida S., Takeda K., Maruoka Y., Sasaki S.;
 RA "Changes in homeobox-containing gene expression during ectopic bone
 RT formation induced by bone morphogenetic protein."
 RL Biochem. Biophys. Res. Commun. 201:980-987(1994).
 DR EMBL: S71286; AAB31006.1;
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox; 1.
 DR Prodom: PD000010; Homeobox; 1.

KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 25 25
 SO SEQUENCE 25 AA: 3111 MW: 009AAB1607BD2203 CRC64;

063990 Length: 25 December 22, 2002 19:30 Type: P Check: 5506 ..

1 LFNNVLRER RLEISKITNL TDRQV

11AA_SEQUENCE 1.0
 ID O68226 PRELIMINARY: PRT: 33 AA.

AC O68226;
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Mszf76 (Fragment).

OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;

RA MEDLINE-98296253; PubMed-9630514;
 RA Agata Y., Matsuda E., Shimizu A.;
 RT "Rapid and efficient cloning of cDNAs encoding Kruppel-like zinc
 RT finger proteins by degenerate PCR.";
 RL Gene 213:55-64(1998).
 DR EMBL: AB010321; BAA31377.1; -;
 DR InterPro: IPR000822; Znf_C2H2.
 DR Pfam: PF00006; Zf-C2H2.1.
 DR ProDom: PD00003; Znf_C2H2.1.
 DR SMART: SM00355; Znf_C2H2.1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 1.
 DR DNA-binding; Metal-binding; zinc-finger.
 FT NON_TER 1 1
 FT NON_TER 33 33
 SQ SEQUENCE 33 AA; 3730 MW; DBD31E5B5C54C538 CRC64;

088226 Length: 33 December 22, 2002 19:30 Type: P Check: 2429 ..

1 IHTEKPYRC AECGKAFTR SNLJKHQTTH TGE

!!AA_SEQUENCE 1.0
 ID Q9JUD0 PRELIMINARY; PRT; 32 AA.
 AC Q9JUD0:
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Protocadherin-T5 (Fragment).
 GN PCDH-T5.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FISCHER344; TISSUE-TESTIS;
 RX MEDLINE-20114371; PubMed-10650949;
 RA Johnson K.J., Patel S.R., Boekelheide K.;
 RT "Multiple cadherin superfamily members with unique expression profiles
 RT are produced in rat testis.";
 RL Endocrinology 141:675-683(2000).
 DR EMBL: AF177699; AAF87074.1; -;
 DR InterPro: IPR002126; Cadherin.
 DR Pfam: PF00028; cadherin.1.
 FT NON_TER 1 1
 FT NON_TER 32 32
 SQ SEQUENCE 32 AA; 3369 MW; A6159FED42844051 CRC64;

Q9JUD0 Length: 32 December 22, 2002 19:30 Type: P Check: 1050 ..

1 ALQAFEFVNG ATRGSPALS SQALVRYVYL DN

!!AA_SEQUENCE 1.0
 ID Q9QX46 PRELIMINARY; PRT; 25 AA.
 AC Q9QX46:
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Amyloid protein (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-129/SVJ; TISSUE-LIVER;
 RX MEDLINE-20102720; PubMed-10636908;
 RA Bergsdtorf C., Paliga K., Kregor S., Masters C.L., Beyreuther K.;
 RT "Identification of cis-elements Regulating Exon 15 Splicing of the
 RT Amyloid Precursor Protein Pre-mRNA.";
 RL J. Biol. Chem. 275:2046-2056(2000).
 DR EMBL: AF199005; AAF20194.1; -;
 DR EMBL: AF199003; AAF20194.1; JOINED.

DR EMBL: AF199004; AAF20194.1; JOINED.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 2608 MW; 276F026D582421D0 CRC64;

Q9QX46 Length: 25 December 22, 2002 19:30 Type: P Check: 4770 ..

1 ENEVEPVDR PAADRGLTTR FGSL

!!AA_SEQUENCE 1.0
 ID Q9ET00 PRELIMINARY; PRT; 20 AA.
 AC Q9ET00:
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE EIRAH (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-129/SVJ;
 RA Green E.D.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF289665; AAF99335.1; -;
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2188 MW; BAFAC0FB953E8P5 CRC64;

Q9ET00 Length: 20 December 22, 2002 19:30 Type: P Check: 5807 ..

1 MADFVYDDR AYSSFGGRC

!!AA_SEQUENCE 1.0
 ID Q9Z115 PRELIMINARY; PRT; 19 AA.
 AC Q9Z115:
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
 DE Calcineurin A alpha (EC 3.1.3.16) (Fragment).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=93111954; PubMed=1335233;
 RA Chang C., Takeuchi T., Mukai H., Shuntoh H., Kuno T., Tanaka C.;
 RT "Molecular cloning and characterization of the promoter region of the
 RT calcineurin A alpha gene.";
 RL Biochem. J. 288:801-805(1992).
 DR EMBL: D10480; BAA01283.1; -;
 KW Hydrolyase.
 FT NON_TER 19 19
 SQ SEQUENCE 19 AA; 2115 MW; F905F8B134CCEC57 CRC64;

Q9Z115 Length: 19 December 22, 2002 19:30 Type: P Check: 4927 ..

1 MSEKADPK LSTTDYVK

!!AA_SEQUENCE 1.0
 ID Q62256 PRELIMINARY; PRT; 18 AA.
 AC Q62256:
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Spontaneous-specific proenkephalin.
 GN PEK-RS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90287163; PubMed=2355920;
 RA Kilpatrick D.L., Zinn S.A., Fitzgerald M., Higuchi H., Sabol S.L.,
 RA Meyerhardt J.;
 RT "Transcription of the rat and mouse proenkephalin genes is initiated
 RT at distinct sites in spermatogenic and somatic cells."
 RL Mol. Cell. Biol. 10:3717-3726(1990).
 DR EMBL: M5181; AAA40127.1; -
 DR MGI: 104628; Penk-ts.
 SQ SEQUENCE 18 AA; 2043 MW; B96E10CC7049FA76 CRC64;

062256 Length: 18 December 22, 2002 19:30 Type: P Check: 3193 ..

1 MSSGKQDSPW EDRIPPGR

IIAA_SEQUENCE 1.0
 ID 063985 PRELIMINARY; PRT; 25 AA.
 AC 063985;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hox1.8 homeobox homolog protein (Fragment).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NX NCBI_TaxID=10118;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94271262; PubMed=7911662;
 RA Iimura T., Oida S., Takeda K., Maruoka Y., Sasaki S.;
 RT "Changes in homeobox-containing gene expression during ectopic bone
 RT formation induced by bone morphogenetic protein."
 RL Biochem. Biophys. Res. Commun. 201:980-987(1994).
 DR EMBL: S12175; AAB31001.1; -
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox; 1.
 DR ProDom: PD000010; Homeobox; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1 1
 SQ SEQUENCE 25 AA; 3134 MW; C54E951607BD27B1 CRC64;

063985 Length: 25 December 22, 2002 19:30 Type: P Check: 5721 ..

1 LENNYLRER RLEISRVHL TDRQV

IIAA_SEQUENCE 1.0
 ID 061159 PRELIMINARY; PRT; 32 AA.
 AC 061159;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE NMR-AG (Fragment).
 CN RRG1 OR NMR-AG.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129SV;
 RX MEDLINE=96315660; PubMed=8754834;
 RA Shago M., Giguere V.;
 RT "Isolation of a novel retinoic acid-responsive gene by selection of
 RT genomic fragments derived from CpG-island-enriched DNA."
 RL Mol. Cell. Biol. 16:4337-4348(1996).
 DR EMBL: U50384; AAB38132.1; -
 DR MGI: 108048; Rrg1.
 FT NON_TER 32 32
 SQ SEQUENCE 32 AA; 3470 MW; 9571FBD08B1FCA84 CRC64;

061159 Length: 32 December 22, 2002 19:30 Type: P Check: 837 ..

1 MAASQDVS FCGVADRAR GSVEYRYDS IK

IIAA_SEQUENCE 1.0
 ID 0923H1 PRELIMINARY; PRT; 40 AA.
 AC 0923H1;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Aistaleless-related homeobox protein Arx (Fragment).
 OS Cricetus griseus (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetus.
 NX NCBI_TaxID=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Onlra R.H., Zhang Y.H., Guo W., Dipple K., Shih S., Doerr J.,
 RA Huang B.-L., Fu L., Abu-Khalil A., Geschwind D., McCabe E.;
 RT "Human ARX gene: genomic characterization and expression."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY038070; AAK93900.1; -
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1 1
 SQ SEQUENCE 40 AA; 3815 MW; 87A852768D0CB6E CRC64;

0923H1 Length: 40 December 22, 2002 19:30 Type: P Check: 2032 ..

1 PTPAVEGAVA SGALADPATA AADRASSIA AXGSRPSMP

IIAA_SEQUENCE 1.0
 ID 099KX5 PRELIMINARY; PRT; 27 AA.
 AC 099KX5;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 2.8 kDa protein.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC003968; AAO3968.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 27 AA; 2832 MW; 4008F368AF868B63 CRC64;

099KX5 Length: 27 December 22, 2002 19:30 Type: P Check: 9106 ..

1 MASLPVGSQ VPADRGYIGP RPAVQEH

IIAA_SEQUENCE 1.0
 ID 090VD0 PRELIMINARY; PRT; 24 AA.
 AC 090VD0;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE vitronectin (fragment).
 OS Cavia (guinea pigs).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriognathi; Cavidae.
 NX NCBI_TaxID=10140;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92207982; PubMed=1372829;
 RA Nakashima N., Miyazaki K., Ishikawa M., Yotoko T., Ogawa H.,
 RA Uchibori H., Matsumoto I., Sano N., Hayashi M.;
 RT "Vitronectin diversity in evolution but uniformity in ligand binding
 RT and size of the core polypeptide."
 RL Biochim. Biophys. Acta 1120:1-10(1992).
 FT NON_TER 1 1

FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2736 MW; 0CD91505B6F22D0A CRC64;

Q9QV00 Length: 24 December 22, 2002 19:30 Type: P Check: 3385 ..

1 XXESKGRXT EGFNADRRKXQ XXEL

!!AA_SEQUENCE 1.0
ID Q9QV89 PRELIMINARY; PRT; 32 AA.
AC Q9QV89;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE SLP-14-FATTY acid-binding protein (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=93191717; PubMed=8447836;
RA Bansal M.P., Medina D.;
RT "Expression of fatty acid-binding proteins in the developing mouse
mammary gland";
RL Biochem. Biophys. Res. Commun. 191:61-69(1993).
DR HSP; P05413; 1HMT.
DR InterPro: IPR000566; Lipocln_cytfabp.
DR Pfam: PF00061; 1fpocalln; 1.
FT NON_TER 1 32
FT NON_TER 32 32
SQ SEQUENCE 32 AA; 3495 MW; 67D23A0146A6AFE CRC64;

Q9QV89 Length: 32 December 22, 2002 19:30 Type: P Check: 581 ..

1 EISFOLGYEF DEVTADRRKXV KSVYTLGGC LV

!!AA_SEQUENCE 1.0
ID Q9QV71 PRELIMINARY; PRT; 17 AA.
AC Q9QV71;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE Lactase-phlorizin hydrolase (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=93293888; PubMed=8514793;
RA Dudley M.A., Hachey D.L., Quaroni A., Hutchens T.W., Nichols B.L.,
RT Rosenberger J., Perkins J.S., Cook G., Reeds P.J.;
RT "In vivo sucrose-isomaltase and lactase-phlorizin hydrolase turnover
in the fed adult rat";
RL J. Biol. Chem. 268:13609-13616(1993).
SQ SEQUENCE 17 AA; 1947 MW; 08941FD316F9692F CRC64;

Q9QV71 Length: 17 December 22, 2002 19:30 Type: P Check: 1363 ..

1 DWDENFIAA GPLTNDL

!!AA_SEQUENCE 1.0
ID Q61461 PRELIMINARY; PRT; 25 AA.
AC Q61461;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Cytochrome P-450b (Phenobarbital-inducible) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=84207435; PubMed=5547088;
RA Stupans I., Ikeda T., Kessler D.J., Nebert D.W.;
RT "Characterization of a cDNA clone for mouse phenobarbital-inducible
RT cytochrome P-450b";
RL DNA 3:129-137(1984).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; K02409; AAA37510.1; -.
DR MGI; MGI:88599; Cyp2b13.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
FT NON_TER 1 25
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 2979 MW; B9DEF7A57355A1B0 CRC64;

Q61461 Length: 25 December 22, 2002 19:30 Type: P Check: 4925 ..

1 SHRLPTLDR SKMPTDAVI HETOR

!!AA_SEQUENCE 1.0
ID Q65747 PRELIMINARY; PRT; 19 AA.
AC Q65747;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Outer coat protein (VP2) (Fragment).
OS Bluetongue virus.
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX NCBI_TaxID=12591;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88265863; PubMed=2838961;
RA Gould A.R., Hyatt A.D., Eaton B.T.;
RT "Morphogenesis of a bluetongue virus variant with an amino acid
RT alteration at a neutralization site in the outer coat protein, VP2";
RL Virology 165:23-32(1988).
DR EMBL; M21355; AAA42845.1; -.
DR InterPro: IPR001742; Orb1_VP2.
DR Pfam: PF00898; Orb1_VP2; 1.
KW Coat protein.
FT NON_TER 1 19
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2608 MW; F0428841265D3061 CRC64;

Q65747 Length: 19 December 22, 2002 19:30 Type: P Check: 4484 ..

1 ERLKFEHRN QRDEDRFY

!!AA_SEQUENCE 1.0
ID Q66538 PRELIMINARY; PRT; 35 AA.
AC Q66538;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE 3' proximal protein (Fragment).
OS Ebola virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae;
OC Ebola-like viruses.
OX NCBI_TaxID=11268;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86124724; PubMed=3946083;
RA Kiley M.P., Willuz J., McCormick J.B., Keene J.D.;
RT "Conservation of the 3' terminal nucleotide sequences of Ebola and
RT Marburg virus";
RL Virology 149:251-254(1986).
DR EMBL; M33062; AAA42976.1; -.
FT NON_TER 35 35
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 4142 MW; F7778E3FC6283092 CRC64;

Q66538 Length: 35 December 22, 2002 19:30 Type: P Check: 8137 ..

```

1 MRKINNELSL KPDRNKLK LLIINHVS EPRHS
!!AA_SEQUENCE 1.0
ID 083622 PRELIMINARY: PRT; 9 AA.
AC 083622;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Hypothetical 1.1 kDa protein (Fragment).
OS Murray valley encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OC NCBI_TaxID=11079;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88118912; PubMed=2828633;
RA Hahn C.S., Hahn Y.S., Rice C.M., Lee E., Dalgarno L., Strauss E.G.,
RA Strauss J.H.;
RT "Conserved elements in the 3' untranslated region of flavivirus RNAs
RT and potential cyclization sequences."
RL J. Mol. Biol. 198;33:41(1987).
DR EMBL; M35172; AAA66627.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1055 MW; FE36D40AAB05A2C1 CRC64;

083622 Length: 9 December 22, 2002 19:30 Type: P Check: 3517 ..

1 THVSEDRVL
!!AA_SEQUENCE 1.0
ID 084254 PRELIMINARY: PRT; 22 AA.
AC 084254;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE X protein (Fragment).
OS Bovine papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OC NCBI_TaxID=10571;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89067912; PubMed=2848926;
RA Stamps A.C., Campo M.S.;
RT "Mapping of two novel transcripts of Bovine papillomavirus type 4."
RL J. Gen. Virol. 69;3033-3045(1988).
DR EMBL; M35264; AAA46926.1; -.
FT NON_TER 1
SQ SEQUENCE 22 AA; 2273 MW; 2801BC23480C9CF9 CRC64;

084254 Length: 22 December 22, 2002 19:30 Type: P Check: 9066 ..

1 NAGPRGCTP EDVADRPDL PE
!!AA_SEQUENCE 1.0
ID 09PWC1 PRELIMINARY: PRT; 39 AA.
AC 09PWC1;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE P55-related MAGUK protein Dlg3 (Fragment).
OS Dlg3.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Koenig C., Yan Y.L., Postlethwait J., Wendler S., Campos-Ortega J.A.;

```

```

RT "A recessive mutation leading to vertebral ankylosis is associated
RT with amino acid exchanges in the zebrafish homologue of the human
RT membrane associated guanylate kinase protein Dlg3."
RL Mech. Dev. 84;1-12(1999).
DR EMBL; AF124436; AAD39393.1; -.
DR InterPro; IPR004172; L27.
DR Pfam; PF02828; L27; 1.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 4477 MW; 6A9D5C428F5A2DBF CRC64;

09PWC1 Length: 39 December 22, 2002 19:30 Type: P Check: 9645 ..

1 LYELSLVLPs QLOPHVESPD DRSELIAMFG ERLSLVYK
!!AA_SEQUENCE 1.0
ID 093438 PRELIMINARY: PRT; 29 AA.
AC 093438;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Lbx1 protein (Fragment).
GN Lbx1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Dietrich S., Schubert F.R., Healy C., Sharpe P.T., Lumsden A.;
RT "Specification of the hypaxial musculature."
RT Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RA Schubert F.R., Dietrich S., Chapman S.C., Lumsden A.;
RT "Expression of the Lbx1 gene in the chick embryo."
RT Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ222831; CAAL1016.1; -.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
FT NON_TER 1
SQ SEQUENCE 29 AA; 3352 MW; CABBC02415F1DAF7 CRC64;

093438 Length: 29 December 22, 2002 19:30 Type: P Check: 3493 ..

1 RFLYKYLSP ADRDQIAQL GLTNQVIT
!!AA_SEQUENCE 1.0
ID 090297 PRELIMINARY: PRT; 27 AA.
AC 090297;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Homeobox (Fragment).
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OC NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=RETINA;
RX MEDLINE=93219357; PubMed=8096640;
RA Levine E.M., Schechter N.;
RT "Homeobox genes are expressed in the retina and brain of adult
RT goldfish."
RT Proc. Natl. Acad. Sci. U.S.A. 90;2729-2733(1993).
DR EMBL; I09693; AAA49178.1; -.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.

```

DR Prodrom: PD000010; Homeobox; 1.
ID 007145 PRELIMINARY; PRT; 27 AA.
AC 007145;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Homeodomain (Fragment).
GN HOMEODOMAIN
OS Petromyzon marinus (See lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=93317669; PubMed=8101001;
RA Pendleton J.W., Nagai B.K., Murtha M.T., Ruddle F.H.;
RT "Expansion of the Hox gene family and the evolution of chordates."
RL Proc. Natl. Acad. Sci. U.S.A. 90:6300-6304(1993).
DR EMBL; L14895; AAA02530.1; -;
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00024; HOMEBOX.
DR Prodrom: PD000010; Homeobox; 1.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 3322 MW; 5529340A951607BD CRC64;

1 LFNMVLTFRER RLEISRSVHL TDRQVKI
!!AA_SEQUENCE 1.0
ID 090298 PRELIMINARY; PRT; 27 AA.
AC 090298;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JUN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Homeobox (Fragment).
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=RETINA;
RC MEDLINE=93219357; PubMed=8096640;
RA Levine E.M., Schechter N.;
RT "Homeobox genes are expressed in the retina and brain of adult goldfish."
RL Proc. Natl. Acad. Sci. U.S.A. 90:2729-2733(1993).
DR EMBL; L09694; AAA49179.1; -;
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00024; HOMEBOX.
DR Prodrom: PD000010; Homeobox; 1.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 3375 MW; 5287254E951607BD CRC64;

090298 Length: 27 December 22, 2002 19:30 Type: P Check: 9642 ..
1 LFNMVLTFRER RLEISRSVHL TDRQVKI
!!AA_SEQUENCE 1.0
ID 007145 PRELIMINARY; PRT; 27 AA.
AC 007145;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Homeodomain (Fragment).
GN HOMEODOMAIN
OS Petromyzon marinus (See lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=93317669; PubMed=8101001;
RA Pendleton J.W., Nagai B.K., Murtha M.T., Ruddle F.H.;
RT "Expansion of the Hox gene family and the evolution of chordates."
RL Proc. Natl. Acad. Sci. U.S.A. 90:6300-6304(1993).
DR EMBL; L14895; AAA02530.1; -;
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00024; HOMEBOX.
DR Prodrom: PD000010; Homeobox; 1.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 3322 MW; 5529340A951607BD CRC64;

007145 Length: 27 December 22, 2002 19:30 Type: P Check: 9552 ..
1 LFNMVLTFRER RLEISRGVNL TDRQVKI

!!AA_SEQUENCE 1.0
ID 007150 PRELIMINARY; PRT; 27 AA.
AC 007150;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Homeodomain (Fragment).
GN HOMEODOMAIN
OS Petromyzon marinus (See lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=93317669; PubMed=8101001;
RA Pendleton J.W., Nagai B.K., Murtha M.T., Ruddle F.H.;
RT "Expansion of the Hox gene family and the evolution of chordates."
RL Proc. Natl. Acad. Sci. U.S.A. 90:6300-6304(1993).
DR EMBL; L14900; AAA02535.1; -;
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00024; HOMEBOX.
DR Prodrom: PD000010; Homeobox; 1.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 3319 MW; 53CD02F1671607BD CRC64;

007150 Length: 27 December 22, 2002 19:30 Type: P Check: 9407 ..
1 LFMSVLTFRER RLEISHLSTL TDRQVKI

!!AA_SEQUENCE 1.0
ID 073591 PRELIMINARY; PRT; 14 AA.
AC 073591;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hox C10 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=WHITE LEGHORN;
RC MEDLINE=98141813; PubMed=9473273;
RA Reale F.V., Jr., Mason K., Hunter A.W., Bothwell M.;
RT "Multiplex display polymerase chain reaction amplifies and resolves related sequences sharing a single moderately conserved domain."
RL Anal. Biochem. 256:158-168(1998).
DR EMBL; U34614; AAC36452.1; -;
FT NON_TER 1
FT NON_TER 14
SQ SEQUENCE 14 AA; 1673 MW; 81258FC9E81FDA45 CRC64;

073591 Length: 14 December 22, 2002 19:30 Type: P Check: 8165 ..

1 RLEISKSINL TDRQ
!!AA_SEQUENCE 1.0
ID P83009 PRELIMINARY; PRT; 13 AA.
AC P83009;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Phospholemman (FYXD domain-containing ion transport regulator 1) (Fragment).
DE Lamna nasus.
OS Lamna nasus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Galeoidea; Lamniformes; Alipodidae;
OC Lamna.
OX NCBI_TaxID=7849;
[1]

```

RP SEQUENCE.
RC TISSUE-RECTAL GLAND:
RA Schuurmans Stekhoven F.M.A.H., Flük G., Wendelaar Bonga S.E.;
RT "Phospholemman in the rectal gland of sharks.";
RL Submitted (JUN-2001) to the SWISS-PROT data bank.
CC -1- FUNCTION: INDUCES A HYPERPOLARIZATION-ACTIVATED CHLORIDE CURRENT
CC WHEN EXPRESSED IN XENOPUS OOCYTES. MAY HAVE A FUNCTIONAL ROLE IN
CC MUSCLE CONTRACTION.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MICROSMAL
CC MEMBRANE.
CC -1- PTM: MAJOR PLASMA MEMBRANE SUBSTRATE FOR CAMP-DEPENDENT PROTEIN
CC KINASE (PK-A) AND PROTEIN KINASE C (PK-C) IN SEVERAL DIFFERENT
CC TISSUES. PHOSPHORYLATED IN RESPONSE TO INSULIN AND ADRENERGIC
CC STIMULATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE FXD FAMILY.
DR InterPro: IPR000272; ATP1G1_PLM_MAT8.
DR PROSITE: PS01310; FXD; PARTIAL.
KW Transmembrane; Phosphorylation; Ionic channel; Ion transport;
KW Microsome.
FT NON_TER 13 13
SQ SEQUENCE 13 AA: 1542 MW: 081373C69724A444 CRC64;

P83009 Length: 13 December 22, 2002 19:30 Type: P Check: 7125 ..

1 VSDVPNNDDR FTY

11AA_SEQUENCE 1.0
ID P83010 PRELIMINARY; PRT; 13 AA.
AC P83010;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Phospholemman (FXD domain-containing ion transport regulator 1)
DE (Fragment).
OS Triakis scyllium (leopard shark) (Triakis scyllia).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeidae; Carcharhiniformes; Triakidae;
OC Triakis.
OX NCBI_TaxID=30494;
RN [1]
RP SEQUENCE.
RC TISSUE-RECTAL GLAND:
RA Schuurmans Stekhoven F.M.A.H., Flük G., Wendelaar Bonga S.E.;
RT "Phospholemman in the rectal gland of sharks.";
RL Submitted (JUN-2001) to the SWISS-PROT data bank.
CC -1- FUNCTION: INDUCES A HYPERPOLARIZATION-ACTIVATED CHLORIDE CURRENT
CC WHEN EXPRESSED IN XENOPUS OOCYTES. MAY HAVE A FUNCTIONAL ROLE IN
CC MUSCLE CONTRACTION.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MICROSMAL
CC MEMBRANE.
CC -1- PTM: MAJOR PLASMA MEMBRANE SUBSTRATE FOR CAMP-DEPENDENT PROTEIN
CC KINASE (PK-A) AND PROTEIN KINASE C (PK-C) IN SEVERAL DIFFERENT
CC TISSUES. PHOSPHORYLATED IN RESPONSE TO INSULIN AND ADRENERGIC
CC STIMULATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE FXD FAMILY.
DR InterPro: IPR000272; ATP1G1_PLM_MAT8.
DR PROSITE: PS01310; FXD; PARTIAL.
KW Transmembrane; Phosphorylation; Ionic channel; Ion transport;
KW Microsome.
FT NON_TER 13 13
SQ SEQUENCE 13 AA: 1497 MW: 1D140C92C1AE1444 CRC64;

P83010 Length: 13 December 22, 2002 19:30 Type: P Check: 6920 ..

1 AGEPPANNEDR FNY

```

```

GN MF-HOXC11A.
OS Oryzias latipes (Medaka fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorphi; Actinopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kurosawa G., Yamada K., Ishiguro H., Hori H.;
RT "Isolation of BAC Clones That Cover Seven Hox Clusters in Medaka
RT Genome.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB055740; BAB62853.1; -.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR ProDom: PD000010; Homeobox; 1.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA: 3368 MW: 548C73FA983E9158 CRC64;

Q90ZG3 Length: 27 December 22, 2002 19:30 Type: P Check: 9516 ..

1 FENVYINKEK RLQLSRLNL TDROYKI

11AA_SEQUENCE 1.0
ID Q9PS38 PRELIMINARY; PRT; 20 AA.
AC Q9PS38;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE PO GLYCOPROTEIN-MYELIN major structural protein (Fragment).
DE Rana catesbeiana (Bull frog).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE.
RX MEDLINE=93002191; PubMed=1382532;
RA Karthigasan J., Bauer T.K., Teplow D.B., Saavedra R.A.;
RA Kirschner D.A.;
RT "Phylogenetically conserved amino acids of MBP and P0 from amphibian
RT myelin.";
RT J. Mol. Neurosci. 3:185-188(1992).
FT NON_TER 1 1
FT NON_TER 20 20
SQ SEQUENCE 20 AA: 2351 MW: AB238ACED7837676 CRC64;

Q9PS38 Length: 20 December 22, 2002 19:30 Type: P Check: 6634 ..

1 IEVYTDREIQ SNVGSVHLX

11AA_SEQUENCE 1.0
ID Q9PS32 PRELIMINARY; PRT; 23 AA.
AC Q9PS32;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Nonmuscle alpha-actinin 115 kDa isoform (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE=93100311; PubMed=1334489;
RA Imanura M., Masaki T.;
RT "A novel nonmuscle alpha-actinin. Purification and characterization of
RT chicken lung alpha-actinin.";
RL J. Biol. Chem. 267:25927-25933(1992).
FT NON_TER 1 1
FT NON_TER 23 23

```

```

SQ SEQUENCE 23 AA; 2823 MW; EBED28437F668363 CRC64;
O9PS32 Length: 23 December 22, 2002 19:30 Type: P Check: 1521 ..

1 LASDLEWIR RTIPWLEDRS POK

11AA_SEQUENCE 1.0
ID O9PRV3 PRELIMINARY; PRT; 24 AA.
AC O9PRV3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE NAF-185 NEUROGLIA-associated protein (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1];
RP MEDLINE=95050502; PubMed=7961672;
RX Solca F.F., Lurie D.I., Diltz C.D., Johnson R.S., Kumar S.,
RA Rubel E.W., Fischer E.H.;
RT "Identification and purification of a chicken brain neuroglia-
associated protein."
RL J. Biol. Chem. 269:27559-27565(1994).
SQ SEQUENCE 24 AA; 2404 MW; 73E0E067F253442 CRC64;

O9PRV3 Length: 24 December 22, 2002 19:30 Type: P Check: 2489 ..

1 DGEDRDAV EEAVLGTGGC RTPK

11AA_SEQUENCE 1.0
ID O9DU42 PRELIMINARY; PRT; 25 AA.
AC O9DU42;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Rev protein (Fragment).
GN REV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1];
RP SEQUENCE FROM N.A.
RX STRAIN=MDU2;
RC MEDLINE=21002573; PubMed=11118069;
RA Motomura K., Kusagawa S., Kato K., Nohomi K., Iwin H.H., Tun K.M.,
RA Thwe M., Oo K.Y., Lwin S., Kyaw O., Zaw M., Nagai Y., Takebe Y.;
RT "Emergence of new forms of human immunodeficiency virus type 1
intersubtype recombinants in central Myanmar."
RL AIDS Res. Hum. Retroviruses 16:1831-1843(2000).
DR EMBL; AB043899; BAB19213.1;
DR Pfam; PF00424; REV; 1.
FT NON_TER 25
SQ SEQUENCE 25 AA; 2964 MW; 0F92F42A1DDE2758 CRC64;

O9DU42 Length: 25 December 22, 2002 19:30 Type: P Check: 5339 ..

1 MAGRESDRDE ELKTVRLIK LIXQS

11AA_SEQUENCE 1.0
ID O70140 PRELIMINARY; PRT; 9 AA.
AC O70140;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Tat Protein (Fragment).
GN TAR.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;

```

```

RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=020;
RX MEDLINE=95194694; PubMed=7888189;
RA Gao F., Yue L., Craig S., Thornton C.L., Robertson D.L.,
RA McCutchan F.E., Brada J.A., Sharp P.M., Hahn B.H.;
RT "Genetic variation of HIV type 1 in four World Health Organization-
sponsored vaccine evaluation sites: generation of functional envelope
(glycoprotein 160) clones representative of sequence subtypes A, B, C,
and E. WHO Network for HIV Isolation and Characterization."
RL AIDS Res. Hum. Retroviruses 10:1359-1368(1994).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN=020;
RX MEDLINE=96190564; PubMed=8627686;
RA Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S.,
RA Karlsson G., Sodroski J., Morgado M., Galvao-Castro B.,
RA von Briesen H., Beddows S., Weber J., Sharp P.M., Shaw G.M.,
RA Hahn B.H.;
RT "Molecular cloning and analysis of functional envelope genes from
human immunodeficiency virus type 1 sequence subtypes A through G. The
WHO and NIAID Networks for HIV Isolation and Characterization."
RL J. Virol. 70:1651-1657(1996).
RN [3];
RP SEQUENCE FROM N.A.
RC STRAIN=020;
RA Allen E.E.;
RL Submitted (APR-1994) to the EMBL/Genbank/DBJ databases.
DR EMBL; U08794; AAB05175.1;
FT NON_TER 1
SQ SEQUENCE 9 AA; 1098 MW; 5B76D40AB1A01A3 CRC64;

O70140 Length: 9 December 22, 2002 19:30 Type: P Check: 3335 ..

1 SKTETDRFD

11AA_SEQUENCE 1.0
ID O25575 PRELIMINARY; PRT; 23 AA.
AC O25575;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein HP0917.
GN HP0917.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=26695; ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA Loftus B., Richardson D., Dodson R., Adams M.D., Hickey E.K.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Kelley J.M.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
pylori."
RL Nature 388:539-547(1997).
DR EMBL; AE000601; AAD07968.1;
DR TIGR; HP0917;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 23 AA; 2717 MW; FC3042379DA2FD44 CRC64;

O25575 Length: 23 December 22, 2002 19:30 Type: P Check: 575 ..

1 MSPITPLRNP LTOEDRFFOE IIA

```

1.0
ID 006283 PRELIMINARY: PRT: 27 AA.
AC 006283;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein RV3599C.
GN RV3599C OR MTC07H7B.23.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98295967; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
Stalson J.E., Taylor K., Whitehead S., Barrell B.G.,
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
DR EMBL: Z95557; CAB08945.1; -.
DR Tuberculists; RV3599C; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 27 AA; 2894 MW; 1746F2AA97B95C6B CRC64;

006283 Length: 27 December 22, 2002 19:30 Type: P Check: 8577 ..

1 MPASSLCTGS PAADRLDATH ERRREVI
1.0
ID 09KT25 PRELIMINARY: PRT: 33 AA.
AC 09KT25;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein VC0735.
GN VC0735.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,
RA McDonald L., Uitterback T., Fleischmann R.D., Niernan W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL: AE004159; AAF93900.1; -.
DR TIGR: VC0735; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 33 AA; 3661 MW; A3AD77F43D6FE40C CRC64;

09KT25 Length: 33 December 22, 2002 19:30 Type: P Check: 3787 ..

1 MPVSSIPCS ISTALADRV VSRVWCFPL LVL
1.0
ID 09KL13 PRELIMINARY: PRT: 32 AA.
AC 09KL13;

DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein VCA0761.
GN VCA0761.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,
RA McDonald L., Uitterback T., Fleischmann R.D., Niernan W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL: AE004404; AAF96659.1; -.
DR TIGR: VCA0761; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 32 AA; 3898 MW; E08F60B7C8C67F83 CRC64;

09KL13 Length: 32 December 22, 2002 19:30 Type: P Check: 1300 ..

1 MPDLRFFSA ICSUDRRKQT NLFSSDLKQO WR
1.0
ID 08X407 PRELIMINARY: PRT: 13 AA.
AC 08X407;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein z0899.
GN z0899.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157: H7 / ED933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobcek E.J., Davis N.W., Llim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Ian J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
DR EMBL: AF005252; AAG55068.1; -.
DR EMBL: AF005252; AAG55068.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 13 AA; 1575 MW; 5127D365A6BBD763 CRC64;

08X407 Length: 13 December 22, 2002 19:30 Type: P Check: 6940 ..

1 MSTDRRKPVML LFH
1.0
ID 08VJ09 PRELIMINARY: PRT: 35 AA.
AC 08VJ09;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein M13536.
GN M13536.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxId=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / OSHKOSH;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains." to the EMBL/GenBank/DBJ databases.
 RL Submitted (APR 2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE007156; AAK4/8/6.1; -.
 DR TIGR; MT3536; -.
 KM Hypochemical protein.
 SQ SEQUENCE 35 AA; 4204 MW; 897851C24908FC03 CRC64;

Q8VJ09 Length: 35 December 22, 2002 19:30 Type: P Check: 8356 ..

1 MKYKLAIDE YDRADTERG AILRENLVS SLTTE

!JAA_SEQUENCE 1.0
 ID Q9HNL3 PRELIMINARY; PRT; 32 AA.
 AC Q9HNL3;
 DT 01-MAR-2001 (TREMREL. 16, Created)
 DT 01-MAR-2001 (TREMREL. 16, Last sequence update)
 DT 01-MAR-2001 (TREMREL. 16, Last annotation update)
 DE Vng2049C.
 GN Vng2049C.
 OS Halobacterium sp. (strain NRC-1).
 OC Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxId=64091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20504483; PubMed=11016950.
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leitner B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenberger T.A., Peck R.F., Pohlischer M., Spudis J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.U., Dennis P.P., Omer A.D.,
 RA Edhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
 RT "Genome sequence of Halobacterium species NRC-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 DR EMBL; AE005097; AAG20207.1; -.
 KM Complete proteome.
 SQ SEQUENCE 32 AA; 3719 MW; D0F967E650A578A CRC64;

Q9HNL3 Length: 32 December 22, 2002 19:30 Type: P Check: 9872 ..

1 MVRVPTDAG RAEMERADR LETTAFFWRL VD

...

+

1

,

:

.

!!AA_SEQUENCE 1.0
F1;A21630 - cytochrome P450b - mouse (fragment)
N;Contains: oxidoreductase (EC 1.-.-.-)
C;Species: Mus musculus (house mouse)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 11-Jun-1999
C;Accession: A21630
R;Stupans, I.; Ikeda, T.; Kessler, D.J.; Nebert, D.W.
DNA 3, 129-137, 1984
A;Title: Characterization of a cDNA clone for mouse phenobarbital-inducible
cytochrome p-450b.
A;Reference number: A21630; MUID:84207435; PMID:6547088
A;Accession: A21630
A;Molecule type: mRNA
A;Residues: 1-25 <STU>
A;Cross-references: GB:K02409; NID:g192896; PIDN:AAA37510.1; PID:g553905
A;Note: the authors translated the codon CTA for residue 5 as Pro, ACC for
residue 7 as Ser and AGT for residue 12 as Thr
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: electron transfer; heme; monooxygenase; oxidoreductase;
transmembrane protein

A21630 Length: 25 December 22, 2002 19:21 Type: P Check: 4925 ..

1 SHRLPTLDDR SKMPYTDAVI HEIQR

!!AA_SEQUENCE 1.0
F1;F42762 - C 3.4.25.1 proteasome endopeptidase complex () subunit 13 - bovine
(fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 03-Jun-2002
C;Accession: F42762
R;Dick, L.R.; Moomaw, C.R.; Pramanik, B.C.; DeMartino, G.N.; Slaughter, C.A.
Biochemistry 31, 7347-7355, 1992
A;Title: Identification and localization of a cysteinyl residue critical for
the trypsin-like catalytic activity of the proteasome.
A;Reference number: A42762; MUID:92378961; PMID:1510924
A;Accession: F42762
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <DIC>
A;Note: sequence extracted from NCBI backbone (NCBIP:112175)
C;Superfamily: multicatalytic endopeptidase complex chain C9
C;Keywords: hydrolase

F42762 Length: 20 December 22, 2002 19:21 Type: P Check: 5572 ..

1 MKGKNXVAIA ADRRFGIQAQ

!!AA_SEQUENCE 1.0
F1;S55729 - orotidine-5'-monophosphate decarboxylase - Aspergillus awamori
(fragments)
C;Species: Aspergillus awamori
C;Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C;Accession: S55729
R;Gouka, R.J.; Hessing, J.G.M.; Stam, H.; Musters, W.; van den Hondel,
C.A.M.J.J.
Curr. Genet. 27, 536-540, 1995
A;Title: A novel strategy for the isolation of defined pyrG mutants and the
development of a site-specific integration system for Aspergillus awamori.
A;Reference number: S55729; MUID:96031709; PMID:7553938
A;Accession: S55729
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-18;19-28 <GOU>
C;Superfamily: orotidine-5'-phosphate decarboxylase; orotidine-5'-phosphate
decarboxylase homology

S55729 Length: 28 December 22, 2002 19:21 Type: P Check: 1645 ..

1 VTVSADVTTT KELLDLADRV MGFVSTRS

!!AA_SEQUENCE 1.0
P1;B44336 - neurotoxin Tx3-2 - spider (Phoneutria nigriventer)
C;Species: Phoneutria nigriventer
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Aug-1998
C;Accession: B44336
R;Cordeiro M do, N.; de Figueiredo, S.G.; Valentim A do, C.; Diniz, C.R.; von
Eickstedt, V.R.; Gilroy, J.; Richardson, M.
Toxicon 31, 35-42, 1993
A;Title: Purification and amino acid sequences of six Tx3 type neurotoxins from
the venom of the Brazilian 'armed' spider Phoneutria nigriventer (Keys.).
A;Reference number: A44336; MUID:93190315; PMID:8446961
A;Accession: B44336
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-34 <COR>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:126773)
C;Superfamily: curtatoxin
C;Keywords: neurotoxin

B44336 Length: 34 December 22, 2002 19:21 Type: P Check: 2946 ..

1 ACAGLYKKCG KGASPCCEDR PCKCDLAMGN CICK

!!AA_SEQUENCE 1.0
F1;B28563 - hemoglobin chain IV - earthworm (Lumbricus terrestris) (fragment)
C;Species: Lumbricus terrestris (common earthworm)
C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 04-Mar-2000
C;Accession: B28563
R;Gotoh, T.; Shishikura, F.; Snow, J.W.; Ereifej, K.I.; Vinogradov, S.N.; Walz,
D.A.
Biochem. J. 241, 441-445, 1987
A;Title: Two globin strains in the giant annelid extracellular haemoglobins.
A;Reference number: A90337; MUID:87241210; PMID:3593201
A;Accession: B28563
A;Molecule type: protein
A;Residues: 1-22 <GOT>
C;Superfamily: globin; globin homology
C;Keywords: oxygen carrier

B28563 Length: 22 December 22, 2002 19:21 Type: P Check: 9362 ..

1 ADEESCCSYE DRREVRHIWD DV

!!AA_SEQUENCE 1.0
F1;S01808 - hemoglobin AII - tube worm (Lamellibrachia sp.) (fragment)
C;Species: Lamellibrachia sp.
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 04-Mar-2000
C;Accession: S01808
R;Suzuki, T.; Takagi, T.; Ohta, S.
Biochem. J. 255, 541-545, 1988
A;Title: N-terminal amino acid sequence of the deep-sea tube worm haemoglobin
remarkably resembles that of annelid haemoglobin.
A;Reference number: S01807; MUID:89076216; PMID:3202832
A;Accession: S01808
A;Molecule type: protein
A;Residues: 1-24 <SUZ>
C;Superfamily: globin; globin homology
C;Keywords: oxygen carrier

S01808 Length: 24 December 22, 2002 19:21 Type: P Check: 3406 ..

1 SSNSCTTEDR REMQLMWANV WSAQ

!!AA_SEQUENCE 1.0
F1;D60894 - gamma crystallin V - bullfrog (fragment)
C;Species: Rana catesbeiana (bullfrog)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Nov-1997
C;Accession: D60894
R;Chiou, S.H.
J. Protein Chem. 7, 527-534, 1988

A;Title: The protein sequence homology of gamma-crystallins among major vertebrate classes and their DNA sequence homology to heat-shock protein genes.
A;Reference number: A60894; MUID:89351593; PMID:3255376
A;Accession: D60894
A;Molecule type: protein
A;Residues: 1-20 <CHI>
C;Superfamily: beta-crystallin
C;Keywords: duplication; eye lens

D60894 Length: 20 December 22, 2002 19:21 Type: P Check: 6311 ..

1 GKIVFYEDRN FQGRSYECSS

!!AA_SEQUENCE 1.0

F1;S04621 - gamma-crystallin II - milk shark (fragment)
C;Species: Rhizoprionodon acutus (milk shark)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Nov-1997
C;Accession: S04621
R;Chiou, S.H.
FEBS Lett. 250, 25-29, 1989
A;Title: Characterization of gamma-crystallins from eye lenses of shark: closer structural similarity to mammalian than other piscine gamma-crystallins?
A;Reference number: S04621; MUID:89290026; PMID:2737298
A;Accession: S04621
A;Molecule type: protein
A;Residues: 1-20 <CHI>
C;Superfamily: beta-crystallin
C;Keywords: duplication

S04621 Length: 20 December 22, 2002 19:21 Type: P Check: 5853 ..

1 GKITFYEDRG FQGHCECSC

!!AA_SEQUENCE 1.0

F1;A61570 - gamma-crystallin (total) - siamang (fragment)
C;Species: Hylobates syndactylus (siamang)
C;Date: 25-Oct-1994 #sequence_revision 06-Jan-1995 #text_change 21-Nov-1997
C;Accession: A61570
R;Chiou, S.H.; Chang, W.P.; Ting, L.M.; Lai, T.A.; Lin, H.K.
Curr. Eye Res. 7, 1017-1022, 1988
A;Title: Biochemical characterization of lens crystallins from three mammalian species.
A;Reference number: A61570; MUID:89152306; PMID:3229121
A;Accession: A61570
A;Molecule type: protein
A;Residues: 1-20 <CHI>
A;Note: 8-Glu, 9-Gln, 10-Ala, 15-Ser, 18-Thr, and 20-Ser were also found
A;Note: this preparation appears to represent the product of several gamma crystallin genes from one specimen
C;Superfamily: beta-crystallin
C;Keywords: duplication; eye lens

A61570 Length: 20 December 22, 2002 19:21 Type: P Check: 6107 ..

1 GKITFYEDRG FQGRHYECTT

!!AA_SEQUENCE 1.0

F1;B61570 - gamma-crystallin (total) - Formosa sika deer (fragment)
C;Species: Cervus nippon taiouanus (Formosa sika deer)
C;Date: 25-Oct-1994 #sequence_revision 06-Jan-1995 #text_change 21-Nov-1997
C;Accession: B61570
R;Chiou, S.H.; Chang, W.P.; Ting, L.M.; Lai, T.A.; Lin, H.K.
Curr. Eye Res. 7, 1017-1022, 1988
A;Title: Biochemical characterization of lens crystallins from three mammalian species.
A;Reference number: A61570; MUID:89152306; PMID:3229121
A;Accession: B61570
A;Molecule type: protein
A;Residues: 1-20 <CHI>
A;Note: 6-Phe, 14-His, and 15-Cys were also found
A;Note: this preparation appears to represent the product of several gamma

crystallin genes from one specimen
C;Superfamily: beta-crystallin
C;Keywords: duplication; eye lens

B61570 Length: 20 December 22, 2002 19:21 Type: P Check: 6068 ..

1 GKITFYEDRG FQGRHYECSS

!!AA SEQUENCE 1.0

F1;A60502 - myonexin - northern leopard frog (fragment)
C;Species: Rana pipiens (northern leopard frog)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999
C;Accession: A60502; A33087
R;Tidball, J.G.
Dev. Biol. 142, 103-114, 1990
A;Title: Myonexin: an 80-kDa glycoprotein that binds fibronectin and is located
at embryonic myotendinous junctions.
A;Reference number: A60502; MUID:91032554; PMID:2146177
A;Accession: A60502
A;Molecule type: protein
A;Residues: 1-25 <TID>
C;Comment: This protein resembles calsequestrin in sequence but differs in
distribution and physical properties.
C;Superfamily: calsequestrin
C;Keywords: fibronectin binding; glycoprotein; skeletal muscle

A60502 Length: 25 December 22, 2002 19:21 Type: P Check: 4514 ..

1 EDGDFDFPEYD GEDRVIFISL KNYKA

!!AA SEQUENCE 1.0

F1;E33208 - calreticulin, uterine - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 31-Jul-1991 #sequence_revision 31-Jul-1991 #text_change 01-Mar-1996
C;Accession: E33208; F33208
R;Milner, R.E.; Baksh, S.; Shemanko, C.; Carpenter, M.R.; Smillie, L.; Vance,
J.E.; Opas, M.; Michalak, M.
J. Biol. Chem. 266, 7155-7165, 1991
A;Title: Calreticulin, and not calsequestrin, is the major calcium binding
protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic reticulum.
A;Reference number: A33208; MUID:91201375; PMID:2016321
A;Accession: E33208
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-29 <MIL>
A;Note: uterine form
A;Accession: F33208
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <MI2>
A;Note: hepatic form
C;Superfamily: calreticulin

E33208 Length: 29 December 22, 2002 19:21 Type: P Check: 2821 ..

1 EPVVYFKEQF LDGDGWTDRW IESKHKSDF

!!AA SEQUENCE 1.0

F1;C33208 - calreticulin, slow twitch skeletal muscle - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 31-Jul-1991 #sequence_revision 31-Jul-1991 #text_change 07-Feb-1997
C;Accession: C33208
R;Milner, R.E.; Baksh, S.; Shemanko, C.; Carpenter, M.R.; Smillie, L.; Vance,
J.E.; Opas, M.; Michalak, M.
J. Biol. Chem. 266, 7155-7165, 1991
A;Title: Calreticulin, and not calsequestrin, is the major calcium binding
protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic reticulum.
A;Reference number: A33208; MUID:91201375; PMID:2016321
A;Accession: C33208
A;Status: preliminary
A;Molecule type: mRNA

A;Residues: 1-29 <MIL>
C;Superfamily: calreticulin
C;Keywords: skeletal muscle

C33208 Length: 29 December 22, 2002 19:21 Type: P Check: 2871 ..

1 EPVVYFQFQF LDGDGWTDRW IESKHKSDF

!!AA_SEQUENCE 1.0

F1;A33434 - calcium-binding protein - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 21-Feb-1990 #sequence_revision 21-Feb-1990 #text_change 16-Feb-1997
C;Accession: A33434
R;Collins, J.H.; Xi, Z.; Alderson-Lang, B.H.; Treves, S.; Volpe, P.
Biochem. Biophys. Res. Commun. 164, 575-579, 1989
A;Title: Sequence homology of a canine brain calcium-binding protein with
calregulin and the human Ro/SS-A antigen.
A;Reference number: A33434; MUID:90026437; PMID:2803321
A;Accession: A33434
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-24 <COL>
C;Superfamily: calreticulin
C;Keywords: calcium binding

A33434 Length: 24 December 22, 2002 19:21 Type: P Check: 2817 ..

1 EPAIYFKEQF LDGDGXTDRX IESK

!!AA_SEQUENCE 1.0

F1;B41481 - virulence-associated protein (virA 5' region) - Salmonella
typhimurium (fragment)
C;Species: Salmonella typhimurium
C;Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 20-Aug-1999
C;Accession: B41481
R;Gulig, P.A.; Chiodo, V.A.
Infect. Immun. 58, 2651-2658, 1990
A;Title: Genetic and DNA sequence analysis of the Salmonella typhimurium
virulence plasmid gene encoding the 28,000-molecular-weight protein.
A;Reference number: A41481; MUID:90316693; PMID:2164511
A;Accession: B41481
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-36 <GUL>
A;Cross-references: GB:M64295; GB:M33927; NID:g154440; PIDN:AAA27272.1;
PID:g154441
C;Superfamily: virulence-associated protein spvB
C;Keywords: DNA binding; transcription regulation

B41481 Length: 36 December 22, 2002 19:21 Type: P Check: 1642 ..

1 GSQDFASQLS KLRLSDDRTA DTNRIKRIIN MRVLNS

!!AA_SEQUENCE 1.0

F1;S07394 - DNA-invertase - phage D108 (fragment)
C;Species: phage D108
A;Note: host Escherichia coli
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Jun-2000
C;Accession: S07394
R;Szatmari, G.B.; Lapointe, M.; DuBow, M.S.
Nucleic Acids Res. 15, 6691-6704, 1987
A;Title: The right end of transposable bacteriophage D108 contains a 520 base
pair protein-encoding sequence not present in bacteriophage Mu.
A;Reference number: S07394; MUID:87316928; PMID:2957646
A;Accession: S07394
A;Molecule type: DNA
A;Residues: 1-22 <SZA>
A;Cross-references: EMBL:X05926; NID:g14918; PIDN:CAA29365.1; PID:g1364174
C;Genetics:
A;Gene: gin
C;Superfamily: transposase repressor

C;Keywords: DNA binding; DNA integration; DNA recombination

S07394 Length: 22 December 22, 2002 19:21 Type: P Check: 9004 ..

1 YKKHPAKRTH IENDDRINQI DR

!!AA_SEQUENCE 1.0

F1;S35552 - vesicle-associated membrane protein 1 - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 10-Dec-1993 #sequence_revision 27-Feb-1997 #text_change 17-Mar-1999

C;Accession: S35552

R;Patarnello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.
Nature 364, 581-582, 1993

A;Title: Neurotransmission and secretion.

A;Reference number: S35552; MUID:93354436; PMID:8350916

A;Accession: S35552

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-32 <PAT>

C;Superfamily: synaptobrevin

S35552 Length: 32 December 22, 2002 19:21 Type: P Check: 9008 ..

1 ERDQKLSELD DRADALQAGA SQFESSAAKL KR

!!AA_SEQUENCE 1.0

F1;S35553 - vesicle-associated membrane protein 1 - chicken (fragment)

C;Species: Gallus gallus (chicken)

C;Date: 20-May-1994 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999

C;Accession: S35553

R;Patarnello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.
Nature 364, 581-582, 1993

A;Title: Neurotransmission and secretion.

A;Reference number: S35552; MUID:93354436; PMID:8350916

A;Accession: S35553

A;Molecule type: DNA

A;Residues: 1-32 <PAT>

C;Superfamily: synaptobrevin

S35553 Length: 32 December 22, 2002 19:21 Type: P Check: 9118 ..

1 ERDQKLSELD DRADALQAGA SVFESSAAKL KR

!!AA_SEQUENCE 1.0

F1;S35555 - vesicle-associated membrane protein 2 - chicken (fragment)

C;Species: Gallus gallus (chicken)

C;Date: 10-Dec-1993 #sequence_revision 26-Jul-1996 #text_change 17-Mar-1999

C;Accession: S35555

R;Patarnello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.
Nature 364, 581-582, 1993

A;Title: Neurotransmission and secretion.

A;Reference number: S35552; MUID:93354436; PMID:8350916

A;Accession: S35555

A;Molecule type: DNA

A;Residues: 1-32 <PAT>

C;Superfamily: synaptobrevin

C;Keywords: membrane trafficking; transmembrane protein

S35555 Length: 32 December 22, 2002 19:21 Type: P Check: 9033 ..

1 ERDQKLSELD DRADALQAGA SQFETSAAKL KR

!!AA_SEQUENCE 1.0

F1;S35554 - vesicle-associated membrane protein 2 - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 10-Dec-1993 #sequence_revision 27-Feb-1997 #text_change 17-Mar-1999

C;Accession: S35554

R;Patarnello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.
Nature 364, 581-582, 1993

A;Title: Neurotransmission and secretion.

A;Reference number: S35552; MUID:93354436; PMID:8350916

A;Accession: S35554
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-32 <PAT>
C;Superfamily: synaptobrevin

S35554 Length: 32 December 22, 2002 19:21 Type: P Check: 9033 ..

1 ERDQKLSELD DRADALQAGA SQFETSAAKL KR

!!AA_SEQUENCE 1.0

F1;E44621 - homeotic protein Hox 10 (clone 10w) - sea lamprey (fragment)
C;Species: Petromyzon marinus (sea lamprey)
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 15-Oct-1999
C;Accession: E44621
R;Pendleton, J.W.; Nagai, B.K.; Murtha, M.T.; Ruddle, F.H.
Proc. Natl. Acad. Sci. U.S.A. 90, 6300-6304, 1993
A;Title: Expansion of the Hox gene family and the evolution of chordates.
A;Reference number: A44616; MUID:93317669; PMID:8101001
A;Accession: E44621
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-27 <PEN>
A;Cross-references: GB:L14895; NID:g290813; PIDN:AAA02530.1; PID:g290814
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation

E44621 Length: 27 December 22, 2002 19:21 Type: P Check: 9552 ..

1 LFNMYLTRER RLEISRGVNL TDRQVKI

!!AA_SEQUENCE 1.0

F1;F44621 - homeotic protein Hox 10 (clone 10x) - sea lamprey (fragment)
C;Species: Petromyzon marinus (sea lamprey)
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 15-Oct-1999
C;Accession: F44621
R;Pendleton, J.W.; Nagai, B.K.; Murtha, M.T.; Ruddle, F.H.
Proc. Natl. Acad. Sci. U.S.A. 90, 6300-6304, 1993
A;Title: Expansion of the Hox gene family and the evolution of chordates.
A;Reference number: A44616; MUID:93317669; PMID:8101001
A;Accession: F44621
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-27 <PEN>
A;Cross-references: GB:L14900; NID:g290819; PIDN:AAA02535.1; PID:g290820
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation

F44621 Length: 27 December 22, 2002 19:21 Type: P Check: 9407 ..

1 LFSMYLTRER RLEISHLLSL TDRQVKI

!!AA_SEQUENCE 1.0

P1;C84355 - hypothetical protein Vng2049c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: C84355
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.;
Lasky, S.R.; Baliga, N.; Thorsson, V.; Sbrojna, J.; Swartzell, S.; Weir, D.;
Hall, J.; Dahl, T.A.; Welte, R.; Goo, Y.A.; Leithauser, B.; Keller, K.; Cruz,
R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablonski, P.E.; Krebs, M.P.;
Angevine, C.M.; Dale, H.; Isenbarger, T.A.; Peck, R.F.; Pohlschrod, M.;
Spudich, J.L.; Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe,
T.M.; Liang, P.; Riley, M.; Hood, L.; DasSarma, S.
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: C84355
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-32 <STO>
A;Cross-references: GB:AE004437; NID:g10581474; PIDN:AAG20207.1; GSPDB:GN00138
C;Genetics:
A;Gene: VNG2049C

C84355 Length: 32 December 22, 2002 19:21 Type: P Check: 9872 ..

1 MVRVPVTDAG RAEMRERADR LETTAAFWRL VD

!!AA_SEQUENCE 1.0

F1;E64634 - hypothetical protein HP0917 - *Helicobacter pylori* (strain 26695)
C;Species: *Helicobacter pylori*
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C;Accession: E64634
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.;
Fleischmann, R.D.; Ketchum, K.A.; Klenk, H.P.; Gill, S.; Dougherty, B.A.;
Nelson, K.; Quackenbush, J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus,
B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, K.;
Fitzgerald, L.M.; Lee, N.; Adams, M.D.; Hickey, E.K.; Berg, D.E.; Gocayne,
J.D.; Utterback, T.R.; Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman,
J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.;
Fraser, C.M.; Venter, J.C.
A;Title: The complete genome sequence of the gastric pathogen *Helicobacter*
pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: E64634
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-23 <TOM>
A;Cross-references: GB:AE000601; GB:AE000511; NID:g2314051; PIDN:AAD07968.1;
PID:g2314059; TIGR:HP0917

E64634 Length: 23 December 22, 2002 19:21 Type: P Check: 575 ..

1 MSPLTPLRNP LTQEDRFFQE IIA

!!AA_SEQUENCE 1.0

F1;PH0858 - MauD protein - *Paracoccus denitrificans* (fragment)
C;Species: *Paracoccus denitrificans*
C;Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 08-Oct-1999
C;Accession: PH0858
R;Chistoserdov, A.Y.; Boyd, J.; Mathews, F.S.; Lidstrom, M.E.
Biochem. Biophys. Res. Commun. 184, 1181-1189, 1992
A;Title: The genetic organization of the mau gene cluster of the facultative
autotroph *Paracoccus denitrificans*.
A;Reference number: PH0856; MUID:92272706; PMID:1590782
A;Accession: PH0858
A;Molecule type: DNA
A;Residues: 1-23 <CHI>
A;Cross-references: GB:M90098; NID:g150580; PIDN:AAA25577.1; PID:g150581
C;Genetics:
A;Gene: mauD

PH0858 Length: 23 December 22, 2002 19:21 Type: P Check: 807 ..

1 LEADRTGFAS LQQYMASRKK QAA

!!AA_SEQUENCE 1.0

F1;A61412 - methane monooxygenase (EC 1.14.13.25) protein Y - *Methylococcus*
capsulatus (fragment)
N;Alternate names: methane hydroxylase protein Y
C;Species: *Methylococcus capsulatus*
C;Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Sep-1994
C;Accession: A61412
R;Rosenzweig, A.C.; Lippard, S.J.
Acc. Chem. Res. 27, 229-236, 1994
A;Title: Determining the structure of a hydroxylase enzyme that catalyzes the
conversion of methane to methanol in methanotrophic bacteria.
A;Reference number: A61412

A;Accession: A61412
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-27 <ROS>
C;Keywords: oxidoreductase

A61412 Length: 27 December 22, 2002 19:21 Type: P Check: 8072 ..

1 WIEDYASRID FKADRDQIVK AVLAGLK

!!AA_SEQUENCE 1.0

P1;A82288 - hypothetical protein VC0735 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)

C;Species: *Vibrio cholerae*

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C;Accession: A82288

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.;
Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill,
S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.;
Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; McDonald, L.;
Utterback, T.; Fleishmann, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.;
Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: A82288

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-33 <HEI>

A;Cross-references: GB:AE004159; GB:AE003852; NID:g9655167; PIDN:AAF93900.1;
GSPDB:GN00126; TIGR:VC0735

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC0735

A;Map position: 1

A82288 Length: 33 December 22, 2002 19:21 Type: P Check: 3787 ..

1 MPVGSSIPCS ISTALADRKV VSRYWCFRTL LVL

!!AA_SEQUENCE 1.0

P1;B82421 - hypothetical protein VCA0761 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)

C;Species: *Vibrio cholerae*

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C;Accession: B82421

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.;
Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill,
S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.;
Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; McDonald, L.;
Utterback, T.; Fleishmann, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.;
Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: B82421

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-32 <HEI>

A;Cross-references: GB:AE004404; GB:AE003853; NID:g9658174; PIDN:AAF96659.1;
GSPDB:GN00127; TIGR:VCA0761

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VCA0761

A;Map position: 2

B82421 Length: 32 December 22, 2002 19:21 Type: P Check: 1300 ..

1 MPDRLLRFSA ICSTDRRKQT NLFSSDLKQQ WR

!!AA_SEQUENCE 1.0
F1;B39089 - hydrogenase (EC 1.18.99.1) 34K chain - Thiocapsa roseopersicina
(fragment)
C;Species: Thiocapsa roseopersicina
C;Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 23-Jun-1993
C;Accession: B39089
R;Kovacs, K.L.; Tigyi, G.; Thanh, L.T.; Lakatos, S.; Kiss, Z.; Bagyinka, C.
J. Biol. Chem. 266, 947-951, 1991
A;Title: Structural rearrangements in active and inactive forms of hydrogenase
from Thiocapsa roseopersicina.
A;Reference number: A39089; MUID:91093297; PMID:1845998
A;Accession: B39089
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <KOV>
C;Keywords: oxidoreductase

B39089 Length: 20 December 22, 2002 19:21 Type: P Check: 6231 ..

1 XXEQADRPSV IGLPFQEXTG

!!AA_SEQUENCE 1.0
P1;H70954 - hypothetical protein Rv3599c - Mycobacterium tuberculosis (strain
H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: H70954
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.;
Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekaia, F.; Badcock, K.;
Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.;
Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.;
Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail,
M.A.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.;
Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: H70954
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-27 <COL>
A;Cross-references: GB:Z95557; GB:AL123456; NID:g3242276; PIDN:CAB08945.1;
PID:e316847; PID:g2113977
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv3599c

H70954 Length: 27 December 22, 2002 19:21 Type: P Check: 8577 ..

1 MPASSLGTGS PAADRLDATH ERRREVI

!!AA_SEQUENCE 1.0
F1;S14161 - probable dTDP-4-keto-6-deoxy-hexose 3,4-isomerase (EC 5.3.99.-) -
Saccharopolyspora erythraea (fragment)
N;Alternate names: 51K protein
C;Species: Saccharopolyspora erythraea
C;Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 21-Nov-1998
C;Accession: S14161
R;Caffrey, P.; Green, B.; Packman, L.C.; Rawlings, B.J.; Staunton, J.; Leadlay,
P.F.
Eur. J. Biochem. 195, 823-830, 1991
A;Title: An acyl-carrier-protein - thioesterase domain from the
6-deoxyerythronolide B synthase of Saccharopolyspora erythraea. High-level
production, purification and characterisation in Escherichia coli.
A;Reference number: S14091; MUID:91153324; PMID:1999200
A;Accession: S14161
A;Molecule type: protein
A;Residues: 1-20 <CAF>
C;Genetics:

A;Gene: eryCII
C;Function:
A;Description: involved in desosamine biosynthesis
C;Keywords: antibiotic biosynthesis; intramolecular oxidoreductase; isomerase

S14161 Length: 20 December 22, 2002 19:21 Type: P Check: 5957 ..

1 MTTTDRAGLG RQLQMIRGLH

!!AA_SEQUENCE 1.0

F1;S66213 - glucose 1-dehydrogenase (EC 1.1.1.47) - Haloferax mediterranei
(fragment)

C;Species: Haloferax mediterranei

C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 30-Oct-1998

C;Accession: S66213

R;Bonete, M.J.; Pire, C.; LLorca, F.I.; Camacho, M.L.

FEBS Lett. 383, 227-229, 1996

A;Title: Glucose dehydrogenase from the halophilic Archaeon Haloferax
mediterranei: enzyme purification, characterisation and N-terminal sequence.

A;Reference number: S66213; MUID:96198607; PMID:8925901

A;Accession: S66213

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-17 <BON>

C;Keywords: oxidoreductase

S66213 Length: 17 December 22, 2002 19:21 Type: P Check: 1672 ..

1 MKAIAVKRGE DRPVVIE

!!AA_SEQUENCE 1.0

P1;S23180 - carboxypeptidase - Sulfolobus solfataricus

C;Species: Sulfolobus solfataricus

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997

C;Accession: S23180

R;Colombo, S.; D'Auria, S.; Fusi, P.; Zecca, L.; Raia, C.A.; Tortora, P.

Eur. J. Biochem. 206, 349-357, 1992

A;Title: Purification and characterization of a thermostable carboxypeptidase
from the extreme thermophilic archaebacterium Sulfolobus solfataricus.

A;Reference number: S23180; MUID:92283259; PMID:1597179

A;Accession: S23180

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-31 <COL>

S23180 Length: 31 December 22, 2002 19:21 Type: P Check: 7956 ..

1 MDLVEKLKND VKEIEDRIIQ IRRKIXENPP L

!!AA_SEQUENCE 1.0

P1;S29326 - hypothetical protein 22, psbA 5'-region - Japanese black pine
chloroplast

C;Species: chloroplast Pinus thunbergiana (Japanese black pine)

C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 18-Aug-2000

C;Accession: S29326; T07427

R;Tsudzuki, J.; Nakashima, K.; Tsudzuki, T.; Hiratsuka, J.; Shibata, M.;

Wakasugi, T.; Sugiura, M.

Mol. Gen. Genet. 232, 206-214, 1992

A;Title: Chloroplast DNA of black pine retains a residual inverted repeat
lacking rRNA genes: nucleotide sequences of trnQ, trnK, psbA, trnI and trnH and
the absence of rps16.

A;Reference number: S20449; MUID:92212283; PMID:1557027

A;Accession: S29326

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-22 <TSU>

A;Cross-references: EMBL:D11467; NID:g344007; PIDN:BAA02023.1; PID:g344011

R;Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugiura, M.

Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994

A;Title: Loss of all ndh genes as determined by sequencing the entire
chloroplast genome of the black pine Pinus thunbergii.

A;Reference number: Z16030; MUID:95024047; PMID:7937893
A;Accession: T07427
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-22 <WAK>
A;Cross-references: EMBL:D17510; NID:g529643; PIDN:BAA04307.1; PID:g1262588
C;Genetics:
A;Genome: chloroplast
C;Keywords: chloroplast

S29326 Length: 22 December 22, 2002 19:21 Type: P Check: 9119 ..

1 MEYLTTDRSI ECGIYLKKIE SI

!!AA_SEQUENCE 1.0

F1;Tl0123 - probable catalase (EC 1.11.1.6) - cucumber (fragment)
C;Species: Cucumis sativus (cucumber)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: Tl0123
R;Toyama, T.; Teramoto, H.; Takeba, G.; Tsuji, H.
Plant Cell Physiol. 36, 1349-1359, 1995
A;Title: Cytokinin induces a rapid decrease in the levels of mRNAs for
catalase, 3-hydroxy-3-methylglutaryl CoA reductase, lectin and other
unidentified proteins in etiolated cotyledons of cucumber.
A;Reference number: Z16946; MUID:96104306; PMID:8564304
A;Accession: Tl0123
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-23 <TOY>
A;Cross-references: EMBL:D63385; NID:g1199475; PIDN:BAA09701.1; PID:g1199476
A;Experimental source: seedling; cotyledons
C;Keywords: oxidoreductase

Tl0123 Length: 23 December 22, 2002 19:21 Type: P Check: 1444 ..

1 NGSQADRSVG QKLAPHLNVR PSI

!!AA_SEQUENCE 1.0

F1;PS0212 - 29K protein 4228 - rice (strain Nihonbare) (fragment)
C;Species: Oryza sativa (rice)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 11-Apr-1995
C;Accession: PS0212
R;Tsugita, A.; Miyatake, N.
submitted to JIPID, April 1993
A;Reference number: PS0208
A;Accession: PS0212
A;Molecule type: protein
A;Residues: 1-15 <TSU>
A;Experimental source: germ
C;Comment: molecular weight 29K, pI 6.1.

PS0212 Length: 15 December 22, 2002 19:21 Type: P Check: 8963 ..

1 SPADDRRDVG DRYAD

!!AA_SEQUENCE 1.0

F1;B60698 - trichocyst protein 27 - Paramecium tetraurelia (fragment)
C;Species: Paramecium tetraurelia
C;Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 07-Dec-1999
C;Accession: B60698
R;Tindall, S.H.; Devito, L.D.; Nelson, D.L.
J. Cell Sci. 92, 441-447, 1989
A;Title: Biochemical characterization of the proteins of Paramecium secretory
granules.
A;Reference number: A60698; MUID:90078398; PMID:2592449
A;Accession: B60698
A;Molecule type: protein
A;Residues: 1-29 <TIN>
C;Comment: This protein was purified from the extruded matrix (contents) of
trichocysts (secretory granules).
C;Genetics:

A;Genetic code: SGC5
C;Keywords: extracellular protein

B60698 Length: 29 December 22, 2002 19:21 Type: P Check: 2526 ..

1 DPLDRLLSTL TDLEDYVAE QKEDDAKNQ

!!AA_SEQUENCE 1.0

F1;S19614 - globin - polychaete (*Eudistylia vancouveri*) (fragment)
N;Alternate names: chlorocruorin
C;Species: *Eudistylia vancouveri*
C;Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 05-Dec-1998
C;Accession: S19614
R;Qabar, A.N.; Stern, M.S.; Walz, D.A.; Chiu, J.T.; Timkovich, R.; Wall, J.S.; Kapp, O.H.; Vinogradov, S.N.
J. Mol. Biol. 222, 1109-1129, 1991
A;Title: Hierarchy of globin complexes. The quaternary structure of the extracellular chlorocruorin of *Eudistylia vancouverii*.
A;Reference number: S19532; MUID:92106333; PMID:1762147
A;Accession: S19614
A;Molecule type: protein
A;Residues: 1-17 <QAB>
A;Experimental source: plume
C;Complex: dodecamers, each consisting of a trimer of tetramers of globin chains; dodecamers are linked into bilayer structure by Ca(2+) and heme-deficient chimeric globin chains
C;Keywords: calcium; dodecamer; heme; homotetramer; oxygen carrier

S19614 Length: 17 December 22, 2002 19:21 Type: P Check: 1820 ..

1 GMKXXSMEDR KTVLADW

!!AA_SEQUENCE 1.0

F1;A60741 - insulin-like growth factor-binding protein He39L - human (fragment)
C;Species: *Homo sapiens* (man)
C;Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 30-Sep-1993
C;Accession: A60741
R;Forbes, B.; Ballard, F.J.; Wallace, J.C.
J. Endocrinol. 126, 497-506, 1990
A;Title: An insulin-like growth factor-binding protein purified from medium conditioned by a human lung fibroblast cell line (He[39]L) has a novel N-terminal sequence.
A;Reference number: A60741; MUID:91011238; PMID:1698907
A;Accession: A60741
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-25 <FOR>

A60741 Length: 25 December 22, 2002 19:21 Type: P Check: 4352 ..

1 LAPGPGQGVQ AGXPGGXVEE EDRGG

!!AA_SEQUENCE 1.0

F1;B28457 - proteoglycan II, bone - human (tentative sequence) (fragment)
C;Species: *Homo sapiens* (man)
C;Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 31-Mar-2000
C;Accession: B28457
R;Fisher, L.W.; Hawkins, G.R.; Tuross, N.; Termine, J.D.
J. Biol. Chem. 262, 9702-9708, 1987
A;Title: Purification and partial characterization of small proteoglycans I and II, bone sialoproteins I and II, and osteonectin from the mineral compartment of developing human bone.
A;Reference number: A92656; MUID:87250639; PMID:3597437
A;Accession: B28457
A;Molecule type: protein
A;Residues: 1-19 <FIS>

B28457 Length: 19 December 22, 2002 19:21 Type: P Check: 4360 ..

1 DEAXGIAPEV PDDRPFEPS

1

!!AA SEQUENCE 1.0

Fl;C54037 - splicing regulatory protein SWAP homolog (alternatively spliced,
clone pFL2) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 28-Feb-1997
C;Accession: C54037
R;Denhez, F.; Lafyatis, R.
J. Biol. Chem. 269, 16170-16179, 1994
A;Title: Conservation of regulated alternative splicing and identification of
functional domains in vertebrate homologs to the Drosophila splicing regulator,
suppressor-of-white-apricot.
A;Reference number: A54037; MUID:94266805; PMID:8206918
A;Accession: C54037
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-29 <DEN>
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBIP:149928)

C54037 Length: 29 December 22, 2002 19:21 Type: P Check: 1565 ..

1 PQLTQEELEA KQAKQKLEDR LAAAAAREKL

!!AA SEQUENCE 1.0

Fl;A48845 - sterol regulatory element 1 binding protein (alternatively spliced,
clone pCY22) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C;Accession: A48845
R;Yokoyama, C.; Wang, X.; Briggs, M.R.; Admon, A.; Wu, J.; Hua, X.; Goldstein,
J.L.; Brown, M.S.
Cell 75, 187-197, 1993
A;Title: SREBP-1, a basic-helix-loop-helix-leucine zipper protein that controls
transcription of the low density lipoprotein receptor gene.
A;Reference number: A48845; MUID:94006541; PMID:8402897
A;Accession: A48845
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-23 <YOK>
A;Cross-references: GB:S66168; NID:g432596; PIDN:AAB28522.1; PID:g432597
A;Experimental source: HeLa S3 cells
A;Note: sequence extracted from NCBI backbone (NCBIN:138566, NCBIN:138568,
NCBIP:138570)

A48845 Length: 23 December 22, 2002 19:21 Type: P Check: 9975 ..

1 RGRANGTDAP RAGADRGAMD CTF

!!AA SEQUENCE 1.0

Fl;S22228 - vitronectin - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 22-Nov-1993 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
C;Accession: S22228
R;Nakashima, N.; Miyazaki, K.; Ishikawa, M.; Yatohgo, T.; Ogawa, H.; Uchibori,
H.; Matsumoto, I.; Seno, N.; Hayashi, M.
Biochim. Biophys. Acta 1120, 1-10, 1992
A;Title: Vitronectin diversity in evolution but uniformity in ligand binding
and size of the core polypeptide.
A;Reference number: S21768; MUID:92207982; PMID:1372829
A;Accession: S22228
A;Molecule type: protein
A;Residues: 1-25 <NAK>

S22228 Length: 25 December 22, 2002 19:21 Type: P Check: 4810 ..

1 AQESXKGRVT EGFNADRKQQ QDELX

!!AA SEQUENCE 1.0

Fl;A26393 - annexin 36K chain - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 02-Jun-1988 #sequence_revision 02-Jun-1988 #text_change 18-Jun-1993

C;Accession: A26393
R;Geisow, M.J.; Fritsche, U.; Hexham, J.M.; Dash, B.; Johnson, T.
Nature 320, 636-638, 1986
A;Title: A consensus amino-acid sequence repeat in Torpedo and mammalian
Ca(2)+-dependent membrane-binding proteins.
A;Reference number: A93379; MUID:86203621; PMID:2422556
A;Accession: A26393
A;Molecule type: protein
A;Residues: 1-16 <GEI>

A26393 Length: 16 December 22, 2002 19:21 Type: P Check: 7 ..

1 MLGLGTDEDR LIEIIL

!!AA SEQUENCE 1.0
P1;S68260 - hypothetical protein gadd7.1 - long-tailed hamster
C;Species: Cricetulus longicaudatus (long-tailed hamster)
C;Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C;Accession: S68260
R;Hollander, M.C.; Alamo, I.; Fornace Jr., A.J.
Nucleic Acids Res. 24, 1589-1593, 1996
A;Title: A novel DNA damage-inducible transcript, gadd7, inhibits cell growth,
but lacks a protein product.
A;Reference number: S68260; MUID:96211359; PMID:8649973
A;Accession: S68260
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-38 <HOL>
A;Cross-references: EMBL:L40430

S68260 Length: 38 December 22, 2002 19:21 Type: P Check: 6375 ..

1 MTPSFSGSSK QLQRNAQMED RGNHPSEFI TGDNLLKT

!!AA SEQUENCE 1.0
P1;A35678 - hypothetical protein (proenkephalin 5' region) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 05-Nov-1999
C;Accession: A35678
R;Kilpatrick, D.L.; Zinn, S.A.; Fitzgerald, M.; Higuchi, H.; Sabol, S.L.;
Meyerhardt, J.
Mol. Cell. Biol. 10, 3717-3726, 1990
A;Title: Transcription of the rat and mouse proenkephalin genes is initiated at
distinct sites in spermatogenic and somatic cells.
A;Reference number: A35678; MUID:90287163; PMID:2355920
A;Accession: A35678
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-18 <KIL>
A;Cross-references: GB:M55181; NID:g201032; PIDN:AAA40127.1; PID:g201033

A35678 Length: 18 December 22, 2002 19:21 Type: P Check: 3193 ..

1 MSSGKQDSPW EDRIPPGR

!!AA SEQUENCE 1.0
F1;PH0780 - T-cell receptor alpha chain (C11) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0780
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility
complex-restricted cytotoxic T lymphocyte clones specific for a Plasmodium
berghei nonapeptide: implications for T cell allelic exclusion and
antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0780
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
A;Cross-references: EMBL:X60879

A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

PH0780 Length: 15 December 22, 2002 19:21 Type: P Check: 8887 ..

1 CALSETGGAD RLTFG

!!AA_SEQUENCE 1.0

F1;PH0794 - T-cell receptor alpha chain (K1 V-alpha-3.pHDS58) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C;Accession: PH0794

R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.

J. Exp. Med. 174, 1371-1383, 1991

A;Title: T cell receptor genes in a series of class I major histocompatibility complex-restricted cytotoxic T lymphocyte clones specific for a Plasmodium berghei nonapeptide: implications for T cell allelic exclusion and antigen-specific repertoire.

A;Reference number: PH0746; MUID:92078846; PMID:1836010

A;Accession: PH0794

A;Molecule type: mRNA

A;Residues: 1-17 <CAS>

A;Cross-references: EMBL:X60899

A;Experimental source: T lymphocyte

C;Keywords: T-cell receptor

PH0794 Length: 17 December 22, 2002 19:21 Type: P Check: 1531 ..

1 CAVSMNEYRG ADRLTFG

!!AA_SEQUENCE 1.0

F1;PH0806 - T-cell receptor alpha chain (PE5.1.1V-alpha-8.F3.3) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C;Accession: PH0806; PH0781

R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.

J. Exp. Med. 174, 1371-1383, 1991

A;Title: T cell receptor genes in a series of class I major histocompatibility complex-restricted cytotoxic T lymphocyte clones specific for a Plasmodium berghei nonapeptide: implications for T cell allelic exclusion and antigen-specific repertoire.

A;Reference number: PH0746; MUID:92078846; PMID:1836010

A;Accession: PH0806

A;Molecule type: mRNA

A;Residues: 1-15 <CA1>

A;Cross-references: EMBL:X60915

A;Experimental source: T lymphocyte

A;Accession: PH0781

A;Molecule type: mRNA

A;Residues: 1-15 <CA2>

A;Cross-references: EMBL:X60880

A;Experimental source: T lymphocyte

C;Keywords: T-cell receptor

PH0806 Length: 15 December 22, 2002 19:21 Type: P Check: 8864 ..

1 CALSDQGGAD RLTFG

!!AA_SEQUENCE 1.0

F1;S03505 - T-cell receptor alpha chain J region (80) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 30-May-1997

C;Accession: S03505

R;Winoto, A.; Mjolsness, S.; Hood, L.

Nature 316, 832-836, 1985

A;Title: Genomic organization of the genes encoding mouse T-cell receptor alpha-chain.

A;Reference number: S03503; MUID:85296332; PMID:2993908

A;Accession: S03505

A;Molecule type: DNA

A;Residues: 1-20 <WIN>
A;Cross-references: EMBL:X02859
A;Note: this sequence was determined from the germline gene
C;Keywords: T-cell receptor

S03505 Length: 20 December 22, 2002 19:21 Type: P Check: 6013 ..

1 NTEGADRLTF GKGTQLIIQP

!!AA_SEQUENCE 1.0
F1;PT0212 - T-cell receptor alpha chain V-J region (4-1-E.2) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 30-May-1997
C;Accession: PT0212
R;Nakano, N.; Kikutani, H.; Nishimoto, H.; Kishimoto, T.
J. Exp. Med. 173, 1091-1097, 1991
A;Title: T cell receptor V gene usage of islet beta cell-reactive T cells is
not restricted in non-obese diabetic mice.
A;Reference number: PT0209; MUID:91217621; PMID:1902501
A;Accession: PT0212
A;Molecule type: mRNA
A;Residues: 1-10 <NAK>
C;Keywords: T-cell receptor

PT0212 Length: 10 December 22, 2002 19:21 Type: P Check: 3993 ..

1 CAVAGGADRL

!!AA_SEQUENCE 1.0
F1;PT0547 - T-cell receptor beta chain V-D-J region (126-1AI) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0547
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N
regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0547
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-8 <FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

PT0547 Length: 8 December 22, 2002 19:21 Type: P Check: 2627 ..

1 ASSDADRG

!!AA_SEQUENCE 1.0
F1;PT0676 - T-cell receptor beta chain V-D-J region (140-1AL) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0676
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N
regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0676
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

PT0676 Length: 7 December 22, 2002 19:21 Type: P Check: 2049 ..

1 ASGEDRG

!!AA_SEQUENCE 1.0

F1;PT0576 - T-cell receptor beta chain V-D-J region (141-1G) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0576
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0576
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: day 19 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

PT0576 Length: 7 December 22, 2002 19:21 Type: P Check: 2172 ..

1 ASSDDRT

!!AA_SEQUENCE 1.0

F1;PT0366 - T-cell receptor beta chain V-J region (6R2) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 30-May-1997
C;Accession: PT0366
R;Lehmann, P.V.; Drexler, K.; Tary-Lehmann, M.; Falcioni, F.; Hurtenbach, U.; Nagy, Z.A.
J. Exp. Med. 173, 333-341, 1991
A;Title: Graft-versus-host resistance induced by class II major histocompatibility complex-specific T cell clones.
A;Reference number: PT0360; MUID:91108330; PMID:1824856
A;Accession: PT0366
A;Molecule type: mRNA
A;Residues: 1-28 <LEH>
C;Keywords: T-cell receptor

PT0366 Length: 28 December 22, 2002 19:21 Type: P Check: 480 ..

1 LYFCASSEDR NNQLRFLERG LDFSVLED

!!AA_SEQUENCE 1.0

F1;A46592 - lactase-phlorizin hydrolase, 200K isoform - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 01-Nov-1996
C;Accession: A46592
R;Dudley, M.A.; Hachey, D.L.; Quaroni, A.; Hutchens, T.W.; Nichols, B.L.; Rosenberger, J.; Perkinson, J.S.; Cook, G.; Reeds, P.J.
J. Biol. Chem. 268, 13609-13616, 1993
A;Title: In vivo sucrose-isomaltase and lactase-phlorizin hydrolase turnover in the fed adult rat.
A;Reference number: A46592; MUID:93293888; PMID:8514793
A;Accession: A46592
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-17 <DUD>
A;Note: sequence extracted from NCBI backbone (NCBIP:134559)
C;Keywords: carbohydrate digestion; intestine

A46592 Length: 17 December 22, 2002 19:21 Type: P Check: 1363 ..

1 DWEDRNFIAA GPLTNDL

!!AA_SEQUENCE 1.0

F1;S78414 - ribosomal protein RL25, mitochondrial [validated] - rat (tentative sequence) (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 25-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 21-Jul-2000
C;Accession: S78414
R;Goldschmidt-Reisin, S.; Graack, H.R.
submitted to the Protein Sequence Database, February 1998
A;Reference number: S78411

A;Accession: S78414
A;Molecule type: protein
A;Residues: 1-26 <GOL>
A;Note: the protein is designated as mitochondrial ribosomal protein L25
C;Keywords: mitochondrion; protein biosynthesis; ribosome

S78414 Length: 26 December 22, 2002 19:21 Type: P Check: 6086 ..

1 ISRRXEKKNK IVYPDQLDGE DRRDAE

!!AA_SEQUENCE 1.0

F1;S22227 - vitronectin - guinea pig (fragment)
C;Species: Cavia porcellus (guinea pig)
C;Date: 22-Nov-1993 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
C;Accession: S22227
R;Nakashima, N.; Miyazaki, K.; Ishikawa, M.; Yatohgo, T.; Ogawa, H.; Uchibori, H.; Matsumoto, I.; Seno, N.; Hayashi, M.
Biochim. Biophys. Acta 1120, 1-10, 1992
A;Title: Vitronectin diversity in evolution but uniformity in ligand binding and size of the core polypeptide.
A;Reference number: S21768; MUID:92207982; PMID:1372829
A;Accession: S22227
A;Molecule type: protein
A;Residues: 1-25 <NAK>

S22227 Length: 25 December 22, 2002 19:21 Type: P Check: 5585 ..

1 XXESXKGRXT EGFNADRKKQ XXELX

!!AA_SEQUENCE 1.0

F1;A54226 - light-harvesting protein B-830 alpha-1 chain - Chromatium purpuratum (fragment)
C;Species: Chromatium purpuratum
C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 05-Jan-1996
C;Accession: A54226
R;Kerfeld, C.A.; Yeates, T.O.; Thornber, J.P.
Biochemistry 33, 2178-2184, 1994
A;Title: Purification and characterization of the peripheral antenna of the purple-sulfur bacterium Chromatium purpuratum: evidence of an unusual pigment-protein composition.
A;Reference number: A54226; MUID:94162224; PMID:8117674
A;Accession: A54226
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-26 <KER>
C;Keywords: antenna complex; light-harvesting polypeptide

A54226 Length: 26 December 22, 2002 19:21 Type: P Check: 6170 ..

1 MKVPVMMADE NAKLNNHEDD RKKFFV

!!AA_SEQUENCE 1.0

F1;B54226 - light-harvesting protein B-830 alpha-2 chain - Chromatium purpuratum (fragment)
C;Species: Chromatium purpuratum
C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 05-Jan-1996
C;Accession: B54226
R;Kerfeld, C.A.; Yeates, T.O.; Thornber, J.P.
Biochemistry 33, 2178-2184, 1994
A;Title: Purification and characterization of the peripheral antenna of the purple-sulfur bacterium Chromatium purpuratum: evidence of an unusual pigment-protein composition.
A;Reference number: A54226; MUID:94162224; PMID:8117674
A;Accession: B54226
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-23 <KER>
C;Keywords: antenna complex; light-harvesting polypeptide

B54226 Length: 23 December 22, 2002 19:21 Type: P Check: 634 ..

1 MQVPVMLADK NAKLNHPEDD RKR

!!AA_SEQUENCE 1.0

F1;T46593 - phytoene dehydrogenase [imported] - Mycobacterium marinum (fragment)
C;Species: Mycobacterium marinum
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 21-Jul-2000
C;Accession: T46593
R;Ramakrishnan, L.; Tran, H.T.; Federspiel, N.A.; Falkow, S.
J. Bacteriol. 179, 5862-5868, 1997
A;Title: A crtB homolog essential for photochromogenicity in Mycobacterium
marinum: isolation, characterization, and gene disruption via homologous
recombination.
A;Reference number: Z23096; MUID:97440138; PMID:9294446
A;Accession: T46593
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-38 <RAM>
A;Cross-references: EMBL:U92075; NID:g1928930; PIDN:AAB71427.1; PID:g1928931
A;Experimental source: strain M
C;Genetics:
A;Gene: crtI

T46593 Length: 38 December 22, 2002 19:21 Type: P Check: 6757 ..

1 VPGVGVPTTL ISGRLAADRI TGNTTRSIRH LDLKAQLS

!!AA_SEQUENCE 1.0

P1;H85575 - hypothetical protein Z0899 [imported] - Escherichia coli (strain
O157:H7, substrain EDL933)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: H85575
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose,
D.J.; Mayhew, G.F.; Evans, P.S.; Gregor, J.; Kirkpatrick, H.A.; Posfai, G.;
Hackett, J.; Klink, S.; Boutin, A.; Shao, Y.; Miller, L.; Grotbeck, E.J.;
Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, J.; Anantharaman,
T.S.; Lin, J.; Yen, G.; Schwartz, D.C.; Welch, R.A.; Blattner, F.R.
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: H85575
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-13 <STO>
A;Cross-references: GB:AE005174; NID:g12513665; PIDN:AAG55068.1; GSPDB:GN00145;
UWGP:Z0899
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z0899

H85575 Length: 13 December 22, 2002 19:21 Type: P Check: 6940 ..

1 MSTDRKPVML LFH

!!AA_SEQUENCE 1.0

F1;PA0041 - plastoquinol-plastocyanin reductase (EC 1.10.99.1) - Arabidopsis
thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Jun-1992 #sequence_revision 06-Jan-1995 #text_change 03-Jun-2002
C;Accession: PA0041
R;Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.
submitted to JIPID, July 1994
A;Description: Separation and characterization of Arabidopsis proteins by
two-dimensional gel electrophoresis.
A;Reference number: PA0001
A;Accession: PA0041
A;Molecule type: protein
A;Residues: 1-15 <KAM>
A;Experimental source: leaf
C;Keywords: oxidoreductase

PA0041 Length: 15 December 22, 2002 19:21 Type: P Check: 9117 ..

1 ASSIPADRVP DMEKR

!!AA_SEQUENCE 1.0

ID AL20_CARMA STANDARD; PRT; 27 AA.
 AC P81823;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 20.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
 OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
 OC Brachyura; Eubrachyura; Portunoidea; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD_RES 27 27 AMIDATION (POTENTIAL).
 SQ SEQUENCE 27 AA; 3152 MW; 597A6901965FE987 CRC64;

AL20_CARMA Length: 27 December 22, 2002 19:30 Type: P Check: 9046 ..

1 GYEDEDERP FYALGLGKRP RTYSFGL

!!AA_SEQUENCE 1.0

ID AMD1_CHICK STANDARD; PRT; 26 AA.
 AC P81073;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE AMP deaminase 1 (EC 3.5.4.6) (Myoadenylate deaminase) (AMP deaminase
 DE isoform M) (Fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Breast;
 RX MEDLINE=97269365; PubMed=9114497;
 RA Chilson O.P., Kelly-Chilson A.E., Siegel N.R.;
 RT "AMP-deaminases from chicken and rabbit muscle: partial primary
 RT sequences of homologous 17-kDa CNBr fragments: autorecognition by
 RT rabbit anti-[chicken AMPD].";
 RL Comp. Biochem. Physiol. 116B:371-377(1997).
 CC -!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY
 CC METABOLISM.
 CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).
 CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.
 DR InterPro; IPR001365; A/AMP_deaminase.
 DR PROSITE; PS00485; A DEAMINASE; PARTIAL.
 KW Hydrolase; Nucleotide metabolism; Multigene family.
 FT NON_TER 1 1
 FT NON_TER 26 26
 SQ SEQUENCE 26 AA; 3195 MW; B03E296D63BB6E75 CRC64;

AMD1_CHICK Length: 26 December 22, 2002 19:30 Type: P Check: 7006 ..

1 MNQKHLRFI KKSRYVDADR VVYDAK

!!AA_SEQUENCE 1.0

ID AMD1_RABIT STANDARD; PRT; 26 AA.
AC P81072;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE AMP deaminase 1 (EC 3.5.4.6) (Myoadenylate deaminase) (AMP deaminase
DE isoform M) (Fragment).
GN AMPD1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE=97269365; PubMed=9114497;
RA Chilson O.P., Kelly-Chilson A.E., Siegel N.R.;
RT "AMP-deaminases from chicken and rabbit muscle: partial primary
RT sequences of homologous 17-kDa CNBr fragments: autorecognition by
RT rabbit anti-[chicken AMPD].";
RL Comp. Biochem. Physiol. 116B:371-377(1997).
CC -!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY
CC METABOLISM.
CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).
CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP
CC DEAMINASE 1 IS THE PREDOMINANT FORM IN SKELETAL MUSCLE; AMP
CC DEAMINASE 2 PREDOMINATES IN SMOOTH MUSCLE, NON-MUSCLE TISSUE,
CC EMBRYONIC MUSCLE AND UNDIFFERENTIATED MYOBLASTS; AMP DEAMINASE 3
CC IS FOUND IN ERYTHROCYTES.
CC -!- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.
DR InterPro; IPR001365; A/AMP_deaminase.
DR PROSITE; PS00485; A_DEAMINASE; PARTIAL.
KW Hydrolase; Nucleotide metabolism; Multigene family.
FT NON_TER 1 1
FT NON_TER 26 26
SQ SEQUENCE 26 AA; 3169 MW; B022467EACBB6E75 CRC64;

AMD1_RABIT Length: 26 December 22, 2002 19:30 Type: P Check: 7826 ..

1 MNQKHLRFI KKS YQVDADR VVYSTK

!!AA_SEQUENCE 1.0

ID DNIV_BPD10 STANDARD; PRT; 22 AA.
AC Q38199;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA-invertase (Fragment).
GN GIN.
OS Bacteriophage D108.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC Mu-like viruses.
OX NCBI_TaxID=10671;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87316928; PubMed=2957646;
RA Szatmari G.B., Lapointe M., Dubow M.S.;
RT "The right end of transposable bacteriophage D108 contains a 520 base
RT pair protein-encoding sequence not present in bacteriophage Mu.";
RL Nucleic Acids Res. 15:6691-6703(1987).
CC -!- FUNCTION: THIS PROTEIN CATALYZES THE INVERSION OF A 3000-BP
CC SEGMENT OF PHAGE DNA. THE INVERSION RESULTS IN A MODIFICATION OF
CC THE 3'END OF THE TAIL FIBER GENE AND ALTERS THE HOST SPECIFICITY.
CC -!- SIMILARITY: BELONGS TO THE SITE-SPECIFIC RECOMBINASE RESOLVASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X05926; CAA29365.1; -.
 DR InterPro; IPR001822; Recombinase.
 DR PROSITE; PS00397; RECOMBINASES_1; PARTIAL.
 DR PROSITE; PS00398; RECOMBINASES_2; PARTIAL.
 KW DNA recombination; DNA integration; DNA-binding; DNA invertase.
 FT NON_TER 1 1
 FT DNA_BIND <1 5 H-T-H MOTIF (PROBABLE).
 SQ SEQUENCE 22 AA; 2748 MW; 097E607032767C38 CRC64;

DNIV_BPD10 Length: 22 December 22, 2002 19:30 Type: P Check: 9004 ..

1 YKKHPAKRTH IENDDRINQI DR

!!AA_SEQUENCE 1.0
 ID FIBB_ANTAM STANDARD; PRT; 21 AA.
 AC P14465;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
 GN FGB.
 OS Antilocapra americana (Pronghorn).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Antilocapridae; Antilocapra.
 OX NCBI_TaxID=9891;
 RN [1]
 RP SEQUENCE.
 RA Mross G.A., Doolittle R.F.;
 RT "Amino acid sequence studies on artiodactyl fibrinopeptides."
 RL Arch. Biochem. Biophys. 122:674-684(1967).
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 DR InterPro; IPR002181; Fibrinogen_C.
 DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KW Blood coagulation; Plasma; Sulfation.
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 6 6 SULFATION.
 FT NON_TER 21 21
 SQ SEQUENCE 21 AA; 2585 MW; FCE6183BE1F31627 CRC64;

FIBB_ANTAM Length: 21 December 22, 2002 19:30 Type: P Check: 7016 ..

1 QPSYDYDEEE DDRAKRLRLDA R

!!AA_SEQUENCE 1.0
 ID FIBB_BISBO STANDARD; PRT; 21 AA.
 AC P14466;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
 GN FGB.
 OS Bison bonasus (European bison).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Bovinae; Bison.
 OX NCBI_TaxID=9902;
 RN [1]

RP SEQUENCE.
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR InterPro; IPR002181; Fibrinogen_C.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW Blood coagulation; Plasma; Sulfation.
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
FT MOD_RES 6 6 SULFATION.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2366 MW; 09EE75AF19E6363D CRC64;

FIBB_BISBO Length: 21 December 22, 2002 19:30 Type: P Check: 7213 ..

1 EFPTDYDEGE DDRPKVGLGA R

!!AA_SEQUENCE 1.0

ID FIBB_BUBBU STANDARD; PRT; 21 AA.
AC P14467;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Bubalus bubalis (Domestic water buffalo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bubalus.
OX NCBI_TaxID=89462;
RN [1]
RP SEQUENCE.
RA Mross G.A., Doolittle R.F.;
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";
RL Arch. Biochem. Biophys. 122:674-684(1967).
RN [2]
RP SEQUENCE.
RC STRAIN=Italian breed;
RX MEDLINE=76040091; PubMed=1180969;
RA Balestrieri C., Colonna G., Irace G.;
RT "Covalent structure of fibrinopeptides from buffaloes breeding in
RT Italy.";
RL Biochim. Biophys. Acta 405:517-521(1975).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR InterPro; IPR002181; Fibrinogen_C.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW Blood coagulation; Plasma; Sulfation.
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 6 6 SULFATION.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2379 MW; 09EE75BE4729163D CRC64;

FIBB_BUBBU Length: 21 December 22, 2002 19:30 Type: P Check: 7185 ..

1 QFPTDYDEGQ DDRPKLGLGA R

!!AA_SEQUENCE 1.0

ID FIBB_CEREL STANDARD; PRT; 21 AA.
AC P14468;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Cervus elaphus (Red deer), and
OS Cervus elaphus nelsoni (American elk).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;
OC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=9860, 9864;
RN [1]
RP SEQUENCE.
RC SPECIES=C.elaphus;
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
RN [2]
RP SEQUENCE.
RC SPECIES=C.e.nelsoni;
RA Mross G.A., Doolittle R.F.;
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";
RL Arch. Biochem. Biophys. 122:674-684(1967).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR InterPro; IPR002181; Fibrinogen_C.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW Blood coagulation; Plasma; Sulfation.
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 6 6 SULFATION.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2558 MW; FCEE745D98931627 CRC64;

FIBB_CEREL Length: 21 December 22, 2002 19:30 Type: P Check: 6821 ..

1 QHSTDYDEEE EDRAKLHLDA R

!!AA_SEQUENCE 1.0

ID FIBB_FELCA STANDARD; PRT; 20 AA.
AC P14469;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE.
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS

CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 DR InterPro; IPR002181; Fibrinogen_C.
 DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KW Blood coagulation; Plasma.
 FT PEPTIDE 1 20 FIBRINOPEPTIDE B.
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2328 MW; A829E393B8F627D0 CRC64;

FIBB_FELCA Length: 20 December 22, 2002 19:30 Type: P Check: 5816 ..

1 IIDYYDEGEE DRDVGVDAR

!!AA_SEQUENCE 1.0

ID FIBB_LAMGL STANDARD; PRT; 19 AA.
 AC P14473;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
 GN FGB.
 OS Lama glama (Llama),
 OS Lama vicugna (Vicugna) (Vicugna vicugna), and
 OS Camelus dromedarius (Dromedary) (Arabian camel).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
 OX NCBI_TaxID=9844, 9843, 9838;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=L.glama;
 RA Blomback B., Blomback M., Grondahl N.J.;
 RT "Studies on fibrinopeptides from mammals.";
 RL Acta Chem. Scand. 19:1789-1791(1965).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=C.dromedarius;
 RX MEDLINE=67209145; PubMed=6033721;
 RA Doolittle R.F., Schubert D., Schwartz S.A.;
 RT "Amino acid sequence studies on artiodactyl fibrinopeptides. I.
 RT Dromedary camel, mule deer, and cape buffalo.";
 RL Arch. Biochem. Biophys. 118:456-467(1967).
 RN [3]
 RP SEQUENCE.
 RC SPECIES=L.vicugna;
 RA Mross G.A., Doolittle R.F.;
 RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";
 RL Arch. Biochem. Biophys. 122:674-684(1967).
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 DR InterPro; IPR002181; Fibrinogen_C.
 DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KW Blood coagulation; Plasma; Sulfation.
 FT PEPTIDE 1 19 FIBRINOPEPTIDE B.
 FT MOD_RES 4 4 SULFATION.
 FT NON_TER 19 19
 SQ SEQUENCE 19 AA; 2295 MW; E7EE6B6100568638 CRC64;

FIBB_LAMGL Length: 19 December 22, 2002 19:30 Type: P Check: 4317 ..

1 ATDYDEEEDD RVKVRDLAR


```

!!AA_SEQUENCE 1.0
ID FIBB MUNMU STANDARD; PRT; 21 AA.
AC P14475;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Muntiacus muntjak (Muntjak).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;
OC Cervidae; Muntiacinae; Muntiacus.
OX NCBI_TaxID=9888;
RN [1]
RP SEQUENCE.
RA Mross G.A., Doolittle R.F.;
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";
RL Arch. Biochem. Biophys. 122:674-684(1967).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR InterPro; IPR002181; Fibrinogen_C.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW Blood coagulation; Plasma; Sulfation.
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 6 6 SULFATION.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2514 MW; FCEE75188F0C1627 CRC64;

```

FIBB_MUNMU Length: 21 December 22, 2002 19:30 Type: P Check: 6963 ..

1 QHSTDYDEVE DDRAKLHLDA R

```

!!AA_SEQUENCE 1.0
ID FIBB ODOHE STANDARD; PRT; 21 AA.
AC P14476;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Odocoileus hemionus (Mule deer) (Black-tailed deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;
OC Cervidae; Odocoileinae; Odocoileus.
OX NCBI_TaxID=9872;
RN [1]
RP SEQUENCE.
RX MEDLINE=67209145; PubMed=6033721;
RA Doolittle R.F., Schubert D., Schwartz S.A.;
RT "Amino acid sequence studies on artiodactyl fibrinopeptides. I.
RT Dromedary camel, mule deer, and cape buffalo.";
RL Arch. Biochem. Biophys. 118:456-467(1967).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR InterPro; IPR002181; Fibrinogen_C.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.

```


KW Blood coagulation; Plasma; Sulfation.
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 6 6 SULFATION.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2496 MW; FCF562C51A0C1627 CRC64;

FIBB_ODOHE Length: 21 December 22, 2002 19:30 Type: P Check: 6856 ..

1 QHLADYDEVDDRAKLHLDA R

!!AA_SEQUENCE 1.0

ID FIBB_RANTA STANDARD; PRT; 21 AA.
AC P14479;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Rangifer tarandus (Reindeer) (Caribou).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;
OC Cervidae; Odocoileinae; Rangifer.
OX NCBI_TaxID=9870;
RN [1]
RP SEQUENCE.
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals."
RL Acta Chem. Scand. 19:1789-1791(1965).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR InterPro; IPR002181; Fibrinogen_C.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW Blood coagulation; Plasma; Sulfation.
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 6 6 SULFATION.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2510 MW; FCF562C45F0C1627 CRC64;

FIBB_RANTA Length: 21 December 22, 2002 19:30 Type: P Check: 6866 ..

1 QHLADYDEVEDRAKLHLDA R

!!AA_SEQUENCE 1.0

ID FIBB_SYNCA STANDARD; PRT; 21 AA.
AC P14481;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Syncerus caffer (Cape buffalo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Syncerus.
OX NCBI_TaxID=9970;
RN [1]
RP SEQUENCE.
RX MEDLINE=67209145; PubMed=6033721;
RA Doolittle R.F., Schubert D., Schwartz S.A.;
RT "Amino acid sequence studies on artiodactyl fibrinopeptides. I.
RT Dromedary camel, mule deer, and cape buffalo."
RL Arch. Biochem. Biophys. 118:456-467(1967).

CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 DR InterPro; IPR002181; Fibrinogen_C.
 DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KW Blood coagulation; Plasma; Sulfation.
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 6 6 SULFATION.
 FT NON_TER 21 21
 SQ SEQUENCE 21 AA; 2353 MW; 09EE75AF033B863D CRC64;

FIBB_SYNCA Length: 21 December 22, 2002 19:30 Type: P Check: 7177 ..

1 QFPTDYDEGE DDRPKSGLGA R

!!AA_SEQUENCE 1.0

ID GLU1_ORENI STANDARD; PRT; 36 AA.
 AC P81026;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Glucagon I.
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
 OC Cichlidae; Oreochromis.
 OX NCBI_TaxID=8128;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95384941; PubMed=7656183;
 RA Nguyen T.M., Wright J.R. Jr., Nielsen P.F., Conlon J.M.;
 RT "Characterization of the pancreatic hormones from the Brockmann body
 RT of the tilapia: implications for islet xenograft studies.";
 RL Comp. Biochem. Physiol. 111C:33-44(1995).
 CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 CC THE BLOOD SUGAR LEVEL.
 CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 DR HSSP; P01274; 1GCN.
 DR InterPro; IPR000532; Glucagon.
 DR Pfam; PF00123; hormone2; 1.
 DR PRINTS; PR00275; GLUCAGON.
 DR SMART; SM00070; GLUCA; 1.
 DR PROSITE; PS00260; GLUCAGON; 1.
 KW Glucagon family; Hormone.
 SQ SEQUENCE 36 AA; 4252 MW; 5AAFBF0254425756 CRC64;

GLU1_ORENI Length: 36 December 22, 2002 19:30 Type: P Check: 382 ..

1 HSEGTFNSDY SKYLEDRKAQ DFVRWLMNNK RSGAAE

!!AA_SEQUENCE 1.0

ID LPGE_ECOLI STANDARD; PRT; 19 AA.
 AC P33236;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gef leader peptide.
 GN GEFL OR B0018.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.

OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92048481; PubMed=1943701;
RA Poulsen L.K., Refn A., Molin S., Andersson P.;
RT "The gef gene from Escherichia coli is regulated at the level of
translation.";
RL Mol. Microbiol. 5:1639-1648(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000112; AAC73129.1; ALT_TERM.
DR PIR; S16473; S16473.
DR EcoGene; EG12074; gefL.
KW Leader peptide; Complete proteome.
SQ SEQUENCE 19 AA; 2259 MW; 19B3EDF371EB0BEB CRC64;

LPGE_ECOLI Length: 19 December 22, 2002 19:30 Type: P Check: 4518 ..

1 MLNTRCVPLT DRKVKEKRA

!!AA_SEQUENCE 1.0
ID MY14_EISFO STANDARD; PRT; 14 AA.
AC P46979;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Myoactive tetradecapeptide (ETP).
OS Eisenia foetida (Common brandling worm) (Common dung-worm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbricina; Lumbricidae; Eisenia.
OX NCBI_TaxID=6396;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Gut;
RX MEDLINE=96087879; PubMed=8532604;
RA Ukena K., Oumi T., Matsushima O., Ikeda T., Fujita T., Minakata H.,
RA Nomoto K.;
RT "A novel gut tetradecapeptide isolated from the earthworm, Eisenia
foetida.";
RL Peptides 16:995-999(1995).
CC -!- FUNCTION: HAS A STIMULATIVE EFFECT ON THE CONTRACTION OF GUT
CC MUSCLES.
CC -!- SIMILARITY: TO INSECTS ALLATOTROPIN.
KW Neuropeptide; Amidation.
FT MOD_RES 14 14 AMIDATION.
SQ SEQUENCE 14 AA; 1478 MW; CC9ABEF941CD91AD CRC64;

MY14_EISFO Length: 14 December 22, 2002 19:30 Type: P Check: 7600 ..

1 GFKDGAADRI SHGF

!!AA_SEQUENCE 1.0
ID MY14_PHEVI STANDARD; PRT; 14 AA.
AC P46980;

DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Myoactive tetradecapeptide (PTP).
 OS Pheretima vittata (Earthworm).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
 OC Lumbricina; Megascolecidae; Pheretima.
 OX NCBI_TaxID=46674;
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RC TISSUE=Gut;
 RX MEDLINE=96087879; PubMed=8532604;
 RA Ukena K., Oumi T., Matsushima O., Ikeda T., Fujita T., Minakata H.,
 RA Nomoto K.;
 RT "A novel gut tetradecapeptide isolated from the earthworm, Eisenia
 RT foetida."
 RL Peptides 16:995-999(1995).
 CC -!- FUNCTION: HAS A STIMULATIVE EFFECT ON THE CONTRACTION OF GUT
 CC MUSCLES.
 CC -!- SIMILARITY: TO INSECTS ALLATOTROPIN.
 KW Neuropeptide; Amidation.
 FT MOD RES 14 14 AMIDATION.
 SQ SEQUENCE 14 AA; 1522 MW; DA40BEE67CCD91AD CRC64;

MY14_PHEVI Length: 14 December 22, 2002 19:30 Type: P Check: 7729 ..

1 GFRDGSADRI SHGF

!!AA_SEQUENCE 1.0

ID T2A PARTE STANDARD; PRT; 23 AA.
 AC Q27173;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Trichocyst matrix protein T2-A (Secretory granule protein T2-A)
 DE (TMP 2-A) (Fragment).
 GN T2A.
 OS Paramecium tetraurelia.
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
 OC Paramecium.
 OX NCBI_TaxID=5888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D4-2;
 RX MEDLINE=96059477; PubMed=7579685;
 RA Madeddu L., Gautier M.-C., Vayssie L., Houari A., Sperling L.;
 RT "A large multigene family codes for the polypeptides of the
 RT crystalline trichocyst matrix in Paramecium."
 RL Mol. Biol. Cell 6:649-659(1995).
 RN [2]
 RP PARTIAL SEQUENCE.
 RC STRAIN=D4-2;
 RX MEDLINE=95119139; PubMed=7819344;
 RA Madeddu L., Gautier M.-C., le Caer J.P., de Loubresse N., Sperling L.;
 RT "Protein processing and morphogenesis of secretory granules in
 RT Paramecium."
 RL Biochimie 76:329-335(1994).
 CC -!- FUNCTION: STRUCTURAL PROTEIN THAT CRYSTALLIZE INSIDE THE
 CC TRICHOCYST MATRIX.
 CC -!- SUBCELLULAR LOCATION: TRICHOCYST. THESE ARE ARCHITECTURALLY
 CC COMPLEX SECRETORY STORAGE GRANULES-DOCKED AT THE PLASMA MEMBRANE,
 CC READY TO RAPIDLY RESPOND TO AN EXOCYTOTIC STIMULUS.
 CC -!- SIMILARITY: BELONGS TO THE TMP FAMILY.
 CC -!- DATABASE: NAME=Protein Spotlight;
 CC NOTE=Issue 3 of October 2000;
 CC WWW="http://www.expasy.org/spotlight/articles/sptlt003.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U27509; AAA92609.1; -.
KW Polyprotein; Structural protein; Multigene family.
FT NON_TER 1 1
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2720 MW; 08972B1F2759BDB7 CRC64;

T2A_PARTE Length: 23 December 22, 2002 19:30 Type: P Check: 946 ..

1 DPLDRLLSTL TDLEDYVAE QKE

!!AA_SEQUENCE 1.0

ID T2B_PARTE STANDARD; PRT; 23 AA.
AC Q27174;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trichocyst matrix protein T2-B (Secretory granule protein T2-B)
DE (TMP 2-B) (Fragment).
GN T2B.
OS Paramecium tetraurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
OC Paramecium.
OX NCBI_TaxID=5888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D4-2;
RX MEDLINE=96059477; PubMed=7579685;
RA Madeddu L., Gautier M.-C., Vayssie L., Houari A., Sperling L.;
RT "A large multigene family codes for the polypeptides of the
RT crystalline trichocyst matrix in Paramecium."
RL Mol. Biol. Cell 6:649-659(1995).
RN [2]
RP PARTIAL SEQUENCE.
RC STRAIN=D4-2;
RX MEDLINE=95119139; PubMed=7819344;
RA Madeddu L., Gautier M.-C., le Caer J.P., de Loubresse N., Sperling L.;
RT "Protein processing and morphogenesis of secretory granules in
RT Paramecium."
RL Biochimie 76:329-335(1994).
CC -!- FUNCTION: STRUCTURAL PROTEIN THAT CRYSTALLIZE INSIDE THE
CC TRICHOCYST MATRIX.
CC -!- SUBCELLULAR LOCATION: TRICHOCYST. THESE ARE ARCHITECTURALLY
CC COMPLEX SECRETORY STORAGE GRANULES-DOCKED AT THE PLASMA MEMBRANE,
CC READY TO RAPIDLY RESPOND TO AN EXOCYTOTIC STIMULUS.
CC -!- SIMILARITY: BELONGS TO THE TMP FAMILY.
CC -!- DATABASE: NAME=Protein Spotlight;
CC NOTE=Issue 3 of October 2000;
CC WWW="<http://www.expasy.org/spotlight/articles/sptlt003.html>".

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U27510; AAA92610.1; -.
KW Polyprotein; Structural protein; Multigene family.
FT NON_TER 1 1
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2706 MW; 184D4B1F2759BDB7 CRC64;

T2B_PARTE Length: 23 December 22, 2002 19:30 Type: P Check: 1016 ..

1 DPLDRLVSTL TDLEDYVAE QKE


```

!!AA_SEQUENCE 1.0
ID   Q9ZG33      PRELIMINARY;          PRT;    17 AA.
AC   Q9ZG33;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   Subtilisin/chymotrypsin inhibitor (Fragment).
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=L2 434B;
RA   Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;
RT   "Gene identification of Chlamydia trachomatis by random DNA
RT   sequencing.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087336; AAD04110.1; -.
FT   NON_TER      1          1
FT   NON_TER     17          17
SQ   SEQUENCE    17 AA;  2043 MW;  4FEE704EE041E120 CRC64;

```

Q9ZG33 Length: 17 December 22, 2002 19:30 Type: P Check: 1658 ..

1 NFLGQPFCTD RKNSFRI

```

!!AA_SEQUENCE 1.0
ID   Q9KIL6      PRELIMINARY;          PRT;    24 AA.
AC   Q9KIL6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   F81r (Fragment).
GN   F81R.
OS   Streptomyces coelicolor A3(2).
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=100226;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M145;
RA   Kormanec J., Sevcikova B., Homerova D.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF230491; AAF82064.1; -.
FT   NON_TER     24          24
SQ   SEQUENCE    24 AA;  2524 MW;  FDA1B812ED67583B CRC64;

```

Q9KIL6 Length: 24 December 22, 2002 19:30 Type: P Check: 2042 ..

1 MAADRTGTDE ATAERALGSR APEF

```

!!AA_SEQUENCE 1.0
ID   Q9X9U4      PRELIMINARY;          PRT;    25 AA.
AC   Q9X9U4;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   ADP glucose pyrophosphorylase (Fragment).
GN   GLGC.
OS   Streptomyces coelicolor.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Bruton C.J.;
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);

```


RX MEDLINE=98062210; PubMed=9401038;
RA Martin M., Schneider D., Bruton C.J., Chater K.F., Hardisson C.;
RT "A glgC Gene Essential Only for the First of Two Spatially Distinct
RT Phases of Glycogen Synthesis in Streptomyces coelicolor A3(2).";
RL J. Bacteriol. 179:7784-7789(1997).
DR EMBL; AJ243803; CAB50742.1; -.
DR InterPro; IPR001825; NTP_transferase.
DR Pfam; PF00483; NTP_transferase; 1.
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 2566 MW; 84B94A728A41D25C CRC64;

Q9X9U4 Length: 25 December 22, 2002 19:30 Type: P Check: 4039 ..

1 MLGIVLAGGE GKRLMPLTAD RAKPA

!!AA_SEQUENCE 1.0

ID Q49137 PRELIMINARY; PRT; 32 AA.
AC Q49137;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MxaS protein (Fragment).
GN MXAS.
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AM1;
RX MEDLINE=96074311; PubMed=7592474;
RA Morris C.J., Kim Y.M., Perkins K.E., Lidstrom M.E.;
RT "Identification and nucleotide sequences of mxaA, mxaC, mxaK, mxaL,
RT and mxaD genes from Methylobacterium extorquens AM1.";
RL J. Bacteriol. 177:6825-6831(1995).
DR EMBL; L41608; AAA85567.1; -.
FT NON_TER 1 1
SQ SEQUENCE 32 AA; 3700 MW; EB03BB62B3C56ED5 CRC64;

Q49137 Length: 32 December 22, 2002 19:30 Type: P Check: 9647 ..

1 ALRRICAPFA RPPFRLADRF DAEALSRHLM TT

!!AA_SEQUENCE 1.0

ID Q53914 PRELIMINARY; PRT; 9 AA.
AC Q53914;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CurD protein (Fragment).
OS Streptomyces cyaneus (Streptomyces curacoi).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1904;
RN [1]
RP SEQUENCE FROM N.A.
RA Bergh S.T., Uhlen M.;
RT "Cloning, analysis and heterologous expression of the polyketides
RT synthesis genes of Streptomyces curacoi.";
RL Submitted (JUN-1990) to the EMBL/GenBank/DDBJ databases.
DR EMBL; M33704; AAA26724.1; -.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1027 MW; 995BDDDDC4140AB1 CRC64;

Q53914 Length: 9 December 22, 2002 19:30 Type: P Check: 3396 ..

1 ITDRRAAQP

!!AA_SEQUENCE 1.0

ID Q50476 PRELIMINARY; PRT; 13 AA.
AC Q50476;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Catalase.
 GN KATG.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L10373;
 RA Cockerill F.R., Uhl J.R., Temesgen Z., Zhang Y., Stockman L.,
 RA Roberts G.D., Williams D.L., Kline B.C.;
 RT "Rapid Identification of a point mutation of the Mycobacterium
 RT tuberculosis catalase-peroxidase (katG) gene associated with isoniazid
 RT resistance.";
 RL Submitted (FEB-1994) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; U06263; AAB59976.1; -.
 SQ SEQUENCE 13 AA; 1564 MW; 2F39A45EFE994777 CRC64;

Q50476 Length: 13 December 22, 2002 19:30 Type: P Check: 7152 ..

1 MPEQHPPITD RSR

!!AA_SEQUENCE 1.0
 ID 005422 PRELIMINARY; PRT; 38 AA.
 AC 005422;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE Phytoene dehydrogenase (Fragment).
 GN CRTI.
 OS Mycobacterium marinum.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1781;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M;
 RX MEDLINE=97440138; PubMed=9294446;
 RA Ramakrishnan L., Tran H.T., Federspiel N.A., Falkow S.;
 RT "A crtB homolog essential for photochromogenicity in Mycobacterium
 RT marinum: isolation, characterization, and gene disruption via
 RT homologous recombination.";
 RL J. Bacteriol. 179:5862-5868(1997).
 DR EMBL; U92075; AAB71427.1; -.
 FT NON_TER 1 1
 SQ SEQUENCE 38 AA; 3986 MW; 6E46332707CCDCAB CRC64;

005422 Length: 38 December 22, 2002 19:30 Type: P Check: 6757 ..

1 VPGVGVPTTL ISGRLAADRI TGNTTRSIRH LDLKAQLS

!!AA_SEQUENCE 1.0
 ID Q93A12 PRELIMINARY; PRT; 22 AA.
 AC Q93A12;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Putative ribitol/glucose oxidoreductase (Fragment).
 GN SDRA2.
 OS Thiobacillus ferrooxidans.
 OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.
 OX NCBI_TaxID=920;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC33020;
 RA Bruscella P., Levican G., Ratouchniak J., Holmes D., Bonnefoy V.;
 RT "A second operon encoding a bcl complex in Acidithiobacillus
 RT ferrooxidans.";

RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ311888; CAC44744.1; -.
FT NON_TER 1 1
SQ SEQUENCE 22 AA; 2464 MW; 3725F8E43BBA75C2 CRC64;

Q93A12 Length: 22 December 22, 2002 19:30 Type: P Check: 9419 ..

1 AQNFINPDLD DRTEKDLATS TV

!!AA_SEQUENCE 1.0
ID Q93A08 PRELIMINARY; PRT; 12 AA.
AC Q93A08;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ResB protein (Fragment).
GN RESB.
OS Thiobacillus ferrooxidans.
OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.
OX NCBI_TaxID=920;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC33020;
RA Levican G., Bruscella P., Guacunano M., Inostroza C., Jedlicki E.,
RA Bonnefoy V., Holmes D.S.;
RT "Characterization of the pet and res operons of Acidithiobacillus
RT ferrooxidans.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ413194; CAC88360.1; -.
FT NON_TER 1 1
SQ SEQUENCE 12 AA; 1405 MW; 886AB7DF1E13240A CRC64;

Q93A08 Length: 12 December 22, 2002 19:30 Type: P Check: 5753 ..

1 QSQDDRKENN DG

!!AA_SEQUENCE 1.0
ID Q03977 PRELIMINARY; PRT; 19 AA.
AC Q03977;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Alpha-acceptor polypeptide M15 (Fragment).
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DH5 ALPHA;
RX MEDLINE=93083990; PubMed=1339377;
RA Prentki P.C.;
RT "Nucleotide sequence of the classical lacZ deletion delta M15.";
RL Gene 122:231-232(1992).
DR EMBL; X58252; CAA41206.1; -.
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2092 MW; 8324315E003AA053 CRC64;

Q03977 Length: 19 December 22, 2002 19:30 Type: P Check: 4923 ..

1 MTMITDSLAV VARTDRPSQ

!!AA_SEQUENCE 1.0
ID Q99094 PRELIMINARY; PRT; 36 AA.
AC Q99094;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-FEB-1997 (TrEMBLrel. 02, Last annotation update)
DE MKAA protein (Fragment).
OS Salmonella typhimurium.
OG Plasmid pYA426.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RA Gulig P.A., Chiodo V.A.;
RL Infect. Immun. 58:2651-2658(1991).
DR EMBL; M64295; AAA27272.1; -.
KW Plasmid.
FT NON_TER 1 1
SQ SEQUENCE 36 AA; 4121 MW; C3D43D1C622EBAB7 CRC64;

Q99094 Length: 36 December 22, 2002 19:30 Type: P Check: 1642 ..

1 GSQDFASQLS KLRLSDDRTA DTNRIKRIIN MRVLNS

!!AA_SEQUENCE 1.0
ID Q9NZ25 PRELIMINARY; PRT; 20 AA.
AC Q9NZ25;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE ATP7B (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wu Z.Y., Wang N., MuRong S.X., Lin M.T., Fang L.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF220215; AAF67661.1; -.
DR InterPro; IPR001757; ATPase_E1-E2.
DR Pfam; PF00122; E1-E2_ATPase; 1.
FT NON_TER 1 1
FT VARIANT 19 19 V -> I.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2282 MW; 00268E2C1E0692E3 CRC64;

Q9NZ25 Length: 20 December 22, 2002 19:30 Type: P Check: 6136 ..

1 APIQQLADRF SGYFVPFIVI

!!AA_SEQUENCE 1.0
ID Q15244 PRELIMINARY; PRT; 40 AA.
AC Q15244;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Phosphoribosylpyrophosphate synthetase isoform I (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RA Ishizuka T.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE=92223087; PubMed=1314091;
RA Ishizuka T., Iizasa T., Taira M., Ishijima S., Sonoda T., Shimada H.,
RA Nagatake N., Tatibana M.;
RT "Promoter regions of the human X-linked housekeeping genes PRPS1 and
RT PRPS2 encoding phosphoribosylpyrophosphate synthetase subunit I and II
RT isoforms.";
RL Biochim. Biophys. Acta 1130:139-148(1992).
DR EMBL; D28133; BAA05675.1; -.
DR HSSP; P14193; 1DKU.

FT NON_TER 40 40
SQ SEQUENCE 40 AA; 4446 MW; 57C2B4011860B098 CRC64;

Q15244 Length: 40 December 22, 2002 19:30 Type: P Check: 2868 ..

1 MPNIKIFSGS SHQDLSQKIA DRLGLELGKV VTKKFSNQET

!!AA_SEQUENCE 1.0

ID Q15245 PRELIMINARY; PRT; 40 AA.
AC Q15245;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Phosphoribosylpyrophosphate synthetase isoform II (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RA Ishizuka T.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE=92223087; PubMed=1314091;
RA Ishizuka T., Iizasa T., Taira M., Ishijima S., Sonoda T., Shimada H.,
RA Nagatake N., Tatibana M.;
RT "Promoter regions of the human X-linked housekeeping genes PRPS1 and
RT PRPS2 encoding phosphoribosylpyrophosphate synthetase subunit I and II
RT isoforms.";
RL Biochim. Biophys. Acta 1130:139-148(1992).
DR EMBL; D28134; BAA05676.1; -.
FT NON_TER 40 40
SQ SEQUENCE 40 AA; 4431 MW; CEC2B31686083EFF CRC64;

Q15245 Length: 40 December 22, 2002 19:30 Type: P Check: 3314 ..

1 MPNIVLFSGS SHQDLSQRVA DRLGLELGKV VTKKFSNQET

!!AA_SEQUENCE 1.0

ID Q9BQV8 PRELIMINARY; PRT; 34 AA.
AC Q9BQV8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE L-isoaspartyl/D-aspartyl methyltransferase (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94354847; PubMed=8074695;
RA Tsai W., Clarke S.;
RT "Amino acid polymorphisms of the human L-isoaspartyl/D-aspartyl
RT methyltransferase involved in protein repair.";
RL Biochem. Biophys. Res. Commun. 203:491-497(1994).
DR EMBL; S73902; AAC60639.2; -.
DR InterPro; IPR000682; Pcmt.
DR Pfam; PF01135; PCMT; 1.
KW Methyltransferase; Transferase.
FT NON_TER 1 1
FT NON_TER 34 34
SQ SEQUENCE 34 AA; 3873 MW; EEB18E55F49BF377 CRC64;

Q9BQV8 Length: 34 December 22, 2002 19:30 Type: P Check: 5960 ..

1 NGIIKTDKVF EVMLATDRSH YAKCNPYMDS PQSI


```

!!AA_SEQUENCE 1.0
ID   Q9UCI1      PRELIMINARY;          PRT;      15 AA.
AC   Q9UCI1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Tropomyosin (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=93195352; PubMed=8450225;
RA   Das K.M., Dasgupta A., Mandal A., Geng X.;
RT   "Autoimmunity to cytoskeletal protein tropomyosin. A clue to the
RT   pathogenetic mechanism for ulcerative colitis.";
RL   J. Immunol. 150:2487-2493(1993).
DR   InterPro; IPR000533; Tropomyosin.
DR   Pfam; PF00261; Tropomyosin; 1.
SQ   SEQUENCE 15 AA; 1802 MW; 7A0993CA5A54254C CRC64;

```

Q9UCI1 Length: 15 December 22, 2002 19:30 Type: P Check: 8961 ..

1 HIAEDADRKY EEVAR

```

!!AA_SEQUENCE 1.0
ID   Q96F68      PRELIMINARY;          PRT;      28 AA.
AC   Q96F68;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Unknown (Protein for IMAGE:4563468) (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC011565; AAH11565.1; -.
FT   NON_TER      1      1
SQ   SEQUENCE 28 AA; 2841 MW; 5799D138245D3951 CRC64;

```

Q96F68 Length: 28 December 22, 2002 19:30 Type: P Check: 588 ..

1 VSQPGSCRHG ADRVGHVGQR AGAGVRPE

```

!!AA_SEQUENCE 1.0
ID   Q9NBB1      PRELIMINARY;          PRT;      26 AA.
AC   Q9NBB1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Stretchin-MLCK (Fragment).
GN   STRN-MLCK OR CG8304 OR CG18255.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Champagne M.B., Edwards K.A., Erickson H.P., Kiehart D.P.;
RT   "Drosophila stretchin-MLCK is a Novel Member of the Titin/Myosin Light
RT   Chain Kinase Family.";
RL   J. Mol. Biol. 0:0-0(2000).
DR   EMBL; AF257309; AAF90127.1; -.
DR   FlyBase; FBgn0013988; Strn-Mlck.

```


FT NON_TER 1 1
SQ SEQUENCE 26 AA; 3004 MW; 2F2CB3A55E7FF033 CRC64;

Q9NBB1 Length: 26 December 22, 2002 19:30 Type: P Check: 6662 ..

1 VVARNNFGTD RIFVTVTIKI PKKKEE

!!AA_SEQUENCE 1.0

ID Q9GU45 PRELIMINARY; PRT; 27 AA.
AC Q9GU45;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Homeobox protein Sycox2 (Fragment).
GN SYCOX2.
OS Sycon raphanus.
OC Eukaryota; Metazoa; Porifera; Calcarea; Calcaronea; Leucosoleniida;
OC Sycettidae.
OX NCBI_TaxID=56443;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20476455; PubMed=11020308;
RA Manuel M., Le Parco Y.;
RT "Homeobox Gene Diversification in the Calcareous Sponge, Sycon
RT raphanus.";
RL Mol. Phylogenet. Evol. 17:97-107(2000).
DR EMBL; AF197140; AAG28510.1; -.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2986 MW; CA58FA11B629E3B7 CRC64;

Q9GU45 Length: 27 December 22, 2002 19:30 Type: P Check: 9325 ..

1 KTSKYLAPTT RQALAARLGL TDRQVKV

!!AA_SEQUENCE 1.0

ID Q9GU44 PRELIMINARY; PRT; 27 AA.
AC Q9GU44;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Homeobox protein Sycox3 (Fragment).
GN SYCOX3.
OS Sycon raphanus.
OC Eukaryota; Metazoa; Porifera; Calcarea; Calcaronea; Leucosoleniida;
OC Sycettidae.
OX NCBI_TaxID=56443;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20476455; PubMed=11020308;
RA Manuel M., Le Parco Y.;
RT "Homeobox Gene Diversification in the Calcareous Sponge, Sycon
RT raphanus.";
RL Mol. Phylogenet. Evol. 17:97-107(2000).
DR EMBL; AF197141; AAG28511.1; -.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 3014 MW; CA58E2803F29E3B7 CRC64;

Q9GU44 Length: 27 December 22, 2002 19:30 Type: P Check: 9332 ..

1 RTSKYLAPTT RQALAARLGL TDRQVKV

!!AA_SEQUENCE 1.0

ID Q9XZZ6 PRELIMINARY; PRT; 23 AA.
AC Q9XZZ6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
 DE Calmodulin (Fragment).
 GN CAM.
 OS Littorina saxatilis.
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Mesogastropoda; Littorinoidea; Littorinidae; Littorina.
 OX NCBI_TaxID=31220;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wilding C.S., Grahame J., Mill P.J.;
 RT "Molecular characterisation of calmodulin intron variation in
 RT Littorina (Gastropoda: Prosobranchia) species.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AJ133335; CAB44224.1; -.
 FT NON_TER 1 1
 FT NON_TER 23 23
 SQ SEQUENCE 23 AA; 2815 MW; 5762B9C4E9886A06 CRC64;

Q9XZZ6 Length: 23 December 22, 2002 19:30 Type: P Check: 1592 ..

1 LQDMINEVDA DRQRDDLPR VPY

!!AA_SEQUENCE 1.0
 ID Q94681 PRELIMINARY; PRT; 25 AA.
 AC Q94681;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE PmHbox1 (Fragment).
 OS Polyandrocarpa misakiensis.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Styelidae; Polyandrocarpa.
 OX NCBI_TaxID=7723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USHIMADO;
 RA Fujiwara S., Kawamura K.;
 RT "Cloning of homeobox-containing genes from the budding ascidian
 RT Polyandrocarpa misakiensis.";
 RL Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; D88274; BAA13570.1; -.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 3074 MW; D7CBE263F4B30A51 CRC64;

Q94681 Length: 25 December 22, 2002 19:30 Type: P Check: 5282 ..

1 HFNQYLCRER RQEVAKAVNL TDRQV

!!AA_SEQUENCE 1.0
 ID Q23749 PRELIMINARY; PRT; 27 AA.
 AC Q23749;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Ovxl ortholog homeobox (Fragment).
 GN CTS-OVX1.
 OS Ctenodrilus serratus.
 OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;
 OC Terebellida; Ctenodrilidae; Ctenodrilus.
 OX NCBI_TaxID=40316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94356262; PubMed=7915607;
 RA Dick M.H., Buss L.W.;
 RT "A PCR-based survey of homeobox genes in Ctenodrilus serratus
 RT (Annelida: Polychaeta).";
 RL Mol. Phylogenet. Evol. 3:146-158(1994).

RN [2]
 RP SEQUENCE FROM N.A.
 RA Dick M.H., Buss L.W.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; U26634; AAC46856.1; -.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2879 MW; 4CCE222303AA0102 CRC64;

Q23749 Length: 27 December 22, 2002 19:30 Type: P Check: 8475 ..

1 PQQEILVTDG TIAHRASPET DRGSGEN

!!AA_SEQUENCE 1.0

ID Q25482 PRELIMINARY; PRT; 27 AA.
 AC Q25482;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Anthox1.Ms (Fragment).
 GN ANTHOX1.MS.
 OS Metridium senile (Brown sea anemone) (Frilled sea anemone).
 OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
 OC Nynanthaeae; Metridiidae; Metridium.
 OX NCBI_TaxID=6116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97435515; PubMed=9290214;
 RA Finnerty J.R., Martindale M.Q.;
 RT "Homeoboxes in sea anemones (Cnidaria:Anthozoa): a PCR-based survey of
 RT Nematostella vectensis and Metridium senile.";
 RL Biol. Bull. 193:62-76(1997).
 DR EMBL; U42727; AAA86626.1; -.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 3314 MW; E6B0D191EE3905CF CRC64;

Q25482 Length: 27 December 22, 2002 19:30 Type: P Check: 9269 ..

1 HFNHFLTKER RSEMATQLNL TDRQVKI

!!AA_SEQUENCE 1.0

ID Q94458 PRELIMINARY; PRT; 27 AA.
 AC Q94458;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE CHV-Hb11 protein (Fragment).
 GN CHV-HB11.
 OS Chaetopterus variopedatus.
 OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;
 OC Spionida; Chaetopteridae; Chaetopterus.
 OX NCBI_TaxID=34590;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Irvine S.M., Warinner S.A., Hunter J.D., Martindale M.Q.;
 RT "A survey of homeobox genes in Chaetopterus variopedatus and analysis
 RT of polychaete homeodomains.";
 RL Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; U68283; AAB16992.1; -.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 3351 MW; 74E5C77ED08CDEE4 CRC64;

Q94458 Length: 27 December 22, 2002 19:30 Type: P Check: 8578 ..

1 NQKKFIEKKD RDRISNEIGL DDRQIKY

!!AA_SEQUENCE 1.0

ID_096892 PRELIMINARY; PRT; 27 AA.
 AC 096892;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Caudal (Fragment).
 GN CAD.
 OS Sacculina carcini.
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Maxillopoda; Cirripedia;
 OC Rhizocephala; Kentrogonida; Sacculinidae; Sacculina.
 OX NCBI_TaxID=51650;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98334561; PubMed=9667986;
 RA Mouchel-Vielh E., Rigolot C., Gibert J.M., Deutsch J.S.;
 RT "Molecules and the body plan: the Hox genes of Cirripedes
 RT (Crustacea).";
 RL Mol. Phylogenet. Evol. 9:382-389(1998).
 DR EMBL; U79471; AAD00342.1; -.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 3289 MW; 1CDC2E4CCA04E00F CRC64;

096892 Length: 27 December 22, 2002 19:30 Type: P Check: 9212 ..

1 RFNNYITIKR KLELSRILGL TDRQVKI

!!AA_SEQUENCE 1.0

ID_P91901 PRELIMINARY; PRT; 27 AA.
 AC P91901;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Homeobox protein BHox26 (Fragment).
 OS Beroe ovata.
 OC Eukaryota; Metazoa; Ctenophora; Cyclocoela; Beroida; Beroidae; Beroe.
 OX NCBI_TaxID=10201;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Witchel H.J., Meech R.W.;
 RT "BHox26 - Beroe Homeobox.";
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U89381; AAB49471.1; -.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 3338 MW; 55318651A81607BD CRC64;

P91901 Length: 27 December 22, 2002 19:30 Type: P Check: 9410 ..

1 LFNMYLTRER RLEISKSINL TDRQVKI

!!AA_SEQUENCE 1.0

ID_O02602 PRELIMINARY; PRT; 27 AA.
 AC O02602; O02535;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Homeobox protein BHox35 (Fragment).
 OS Beroe ovata.
 OC Eukaryota; Metazoa; Ctenophora; Cyclocoela; Beroida; Beroidae; Beroe.
 OX NCBI_TaxID=10201;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Witchel H.J., Meech R.W.;
 RT "BHox35 - Beroe Homeobox.";
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U89383; AAB49473.1; -.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 3322 MW; 5529340A951607BD CRC64;

002602 Length: 27 December 22, 2002 19:30 Type: P Check: 9552 ..

1 LFNMYLTRER RLEISRGVNL TDRQVKI

!!AA_SEQUENCE 1.0
 ID O18615 PRELIMINARY; PRT; 34 AA.
 AC O18615;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE FTZ-F1 (Fragment).
 OS Artemia salina (Brine shrimp).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
 OC Artemiidae; Artemia.
 OX NCBI_TaxID=85549;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97338099; PubMed=9192646;
 RA Escriva H., Safi R., Hanni C., Langlois M.C., Saumitou-Laprade P.,
 RA Stehelin D., Capron A., Pierce R., Laudet V.;
 RT "Ligand binding was acquired during evolution of nuclear receptors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:6803-6808(1997).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 DR EMBL; U93418; AAC83395.1; -.
 DR InterPro; IPR001628; Znf_C4steroid.
 DR Pfam; PF00105; zf-C4; 1.
 DR ProDom; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00399; Znf_C4; 1.
 KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
 KW Zinc-finger.
 FT NON_TER 1 1
 FT NON_TER 34 34
 SQ SEQUENCE 34 AA; 4132 MW; 1AE9E6A0D63D4AEF CRC64;

O18615 Length: 34 December 22, 2002 19:30 Type: P Check: 5276 ..

1 FKRTVQNKV YTCMADRSCH IDKSQRKRCF PCRF

!!AA_SEQUENCE 1.0
 ID Q9V973 PRELIMINARY; PRT; 17 AA.
 AC Q9V973;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE CG12485 protein.
 GN CG12485.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003791; AAF57423.1; -.
 DR FlyBase; FBgn0034547; CG12485.
 SQ SEQUENCE 17 AA; 1907 MW; 47D598D29F4860DE CRC64;

Q9V973 Length: 17 December 22, 2002 19:30 Type: P Check: 1645 ..

1 MELSSTDGRP HHAQQGR

!!AA_SEQUENCE 1.0
 ID Q9TWR1 PRELIMINARY; PRT; 21 AA.
 AC Q9TWR1;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE Cysteine proteinase (Fragment).
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Bombycoidea; Bombycidae; Bombyx.
 OX NCBI_TaxID=7091;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94086476; PubMed=8262908;
 RA Takahashi S.Y., Yamamoto Y., Shionoya Y., Kageyama T.;
 RT "Cysteine proteinase from the eggs of the silkmoth, *Bombyx mori*:
 RT identification of a latent enzyme and characterization of activation
 RT and proteolytic processing in vivo and in vitro.";
 RL J. Biochem. 114:267-272(1993).
 SQ SEQUENCE 21 AA; 2379 MW; 9232784B6D6AE4B1 CRC64;

Q9TWR1 Length: 21 December 22, 2002 19:30 Type: P Check: 8106 ..

1 XPEQVDDRKH GAVTDDKXXQ X

!!AA_SEQUENCE 1.0
 ID O62575 PRELIMINARY; PRT; 26 AA.


AC O62575;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hox protein MYX4 (Fragment).
GN HOX OR MYX4.
OS Tetracapsula bryozoides.
OC Eukaryota; Metazoa; Myxozoa; Myxozoa incertae sedis; Tetracapsula.
OX NCBI_TaxID=75694;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98196665; PubMed=9537319;
RA Anderson C.L., Canning E.U., Okamura B.;
RT "A triploblast origin for Myxozoa?";
RL Nature 392:346-347(1998).
DR EMBL; AJ005124; CAA06387.1; -.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1 1
FT NON_TER 26 26
SQ SEQUENCE 26 AA; 3212 MW; E38651AD1657BD2C CRC64;

O62575 Length: 26 December 22, 2002 19:30 Type: P Check: 7242 ..

1 LFNMYLSRER RLEISKSIDL TDRQVK

!!AA_SEQUENCE 1.0

ID Q61676 PRELIMINARY; PRT; 27 AA.
AC Q61676;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE HOX11-D125 protein (D125) (Fragment).
GN HOX11-D125 OR D125.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93281593; PubMed=8099440;
RA Dear T.N., Sanchez-Garcia I., Rabbitts T.H.;
RT "The HOX11 gene encodes a DNA-binding nuclear transcription factor
RT belonging to a distinct family of homeobox genes."
RL Proc. Natl. Acad. Sci. U.S.A. 90:4431-4435(1993).
DR EMBL; L08618; AAA28612.1; -.
DR FlyBase; FBgn0010394; Hox11-D125.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2937 MW; CABBC2F876F2B3A3 CRC64;



Q61676 Length: 27 December 22, 2002 19:30 Type: P Check: 8698 ..

1 LYQKYLSPAD RDEIAASLGL SNAQVIT

!!AA_SEQUENCE 1.0

ID Q9MZW8 PRELIMINARY; PRT; 29 AA.
AC Q9MZW8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Von Willebrand factor (Fragment).
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.

OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20072937; PubMed=10603266;
RA Chaves R., Sampaio I., Schneider M.P., Schneider H., Page S.L.,
RA Goodman M.;
RT "The place of Callimico goeldii in the Callitrichine phylogenetic
RT tree: evidence from von Willebrand factor gene intron II sequences.";
RL Mol. Phylogenet. Evol. 13:392-404(1999).
DR EMBL; AF092833; AAF77601.1; -.
FT NON_TER 1 1
FT NON_TER 29 29
SQ SEQUENCE 29 AA; 3227 MW; 60F1190C8F227D03 CRC64;

Q9MZW8 Length: 29 December 22, 2002 19:30 Type: P Check: 2931 ..

1 GRDCQDHSFS IVIETVQCAD DRDAVCTRS

!!AA_SEQUENCE 1.0
ID Q29394 PRELIMINARY; PRT; 23 AA.
AC Q29394;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Growth hormone receptor (Fragment).
GN GHR.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97049323; PubMed=8894053;
RA Venta P.J., Brouillette J.A., Yuzbasiyan-Gurkan V., Brewer G.J.;
RT "Gene-specific universal mammalian sequence-tagged sites: application
RT to the canine genome.";
RL Biochem. Genet. 34:321-341(1996).
DR EMBL; L77673; AAA97423.1; -.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2620 MW; 85442F0EC46D5D08 CRC64;

Q29394 Length: 23 December 22, 2002 19:30 Type: P Check: 807 ..

1 DDLDEKTEGS DTDRLLSNDH EKS

!!AA_SEQUENCE 1.0
ID Q9TRS7 PRELIMINARY; PRT; 24 AA.
AC Q9TRS7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Vitronectin (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RX MEDLINE=92207982; PubMed=1372829;
RA Nakashima N., Miyazaki K., Ishikawa M., Yatohgo T., Ogawa H.,
RA Uchibori H., Matsumoto I., Seno N., Hayashi M.;
RT "Vitronectin diversity in evolution but uniformity in ligand binding
RT and size of the core polypeptide.";
RL Biochim. Biophys. Acta 1120:1-10(1992).
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1 1
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2745 MW; 94F0054986FE1215 CRC64;

Q9TRS7 Length: 24 December 22, 2002 19:30 Type: P Check: 2610 ..

1 AQESXKGRVT EGFNADRKQQ QDEL

!!AA_SEQUENCE 1.0

ID Q9TQQ9 PRELIMINARY; PRT; 15 AA.
AC Q9TQQ9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE Glutamate dehydrogenase isoform I (EC 1.4.1.2) (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=96061967; PubMed=7588764;
RA Cho S.W., Lee J., Choi S.Y.;
RT "Two soluble forms of glutamate dehydrogenase isoproteins from bovine
RT brain.";
RL Eur. J. Biochem. 233:340-346(1995).
RN [2]
RP SEQUENCE.
RX MEDLINE=96043916; PubMed=7581004;
RA Lee J., Kim S.W., Cho S.W.;
RT "A novel glutamate dehydrogenase from bovine brain: purification and
RT characterization.";
RL Biochem. Mol. Biol. Int. 36:1087-1096(1995).
SQ SEQUENCE 15 AA; 1754 MW; 65F7CD91023AEBA CRC64;

Q9TQQ9 Length: 15 December 22, 2002 19:30 Type: P Check: 8644 ..

1 EEAADREDD PNFFK

!!AA_SEQUENCE 1.0

ID Q9TR40 PRELIMINARY; PRT; 15 AA.
AC Q9TR40;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE Glutamate dehydrogenase isoform II (EC 1.4.1.2) (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=96061967; PubMed=7588764;
RA Cho S.W., Lee J., Choi S.Y.;
RT "Two soluble forms of glutamate dehydrogenase isoproteins from bovine
RT brain.";
RL Eur. J. Biochem. 233:340-346(1995).
SQ SEQUENCE 15 AA; 1724 MW; 65F7CD91023AF925 CRC64;

Q9TR40 Length: 15 December 22, 2002 19:30 Type: P Check: 8661 ..

1 VEAAADREDD PNFFK

!!AA_SEQUENCE 1.0

ID Q9TRF2 PRELIMINARY; PRT; 33 AA.
AC Q9TRF2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE VAMP/synaptobrevin-2 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=93374072; PubMed=8365494;
RA Horikawa H.P., Saisu H., Ishizuka T., Sekine Y., Tsugita A., Odani S.,
RA Abe T.;
RL FEBS Lett. 330:236-240(1993).
DR InterPro; IPR001388; Synaptobrevin.
DR Pfam; PF00957; synaptobrevin; 1.
DR ProDom; PD001229; Synaptobrevin; 1.
DR PROSITE; PS00417; SYNAPTOBREVIN; 1.
SQ SEQUENCE 33 AA; 3672 MW; 584386688353EC55 CRC64;

Q9TRF2 Length: 33 December 22, 2002 19:30 Type: P Check: 1454 ..

1 XNVNDKVLER DQKLSELDDR ADALQAGASQ FET

!!AA_SEQUENCE 1.0

ID Q37112 PRELIMINARY; PRT; 22 AA.
AC Q37112;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE ORF22.
OS Pinus thunbergii (Green pine) (Japanese black pine).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=3350;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92212283; PubMed=1557027;
RA Tsudzuki J., Nakashima K., Tsudzuki T., Hiratsuka J., Shibata M.,
RA Wakasugi T., Sugiura M.;
RT "Chloroplast DNA of black pine retains a residual inverted repeat
RT lacking rRNA genes: nucleotide sequences of trnQ, trnK, psbA, trnI and
RT trnH and the absence of rps16.";
RL Mol. Gen. Genet. 232:206-214(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95094312; PubMed=8001170;
RA Tsudzuki J., Ito S., Tsudzuki T., Wakasugi T., Sugiura M.;
RT "A new gene encoding tRNA(Pro) (GGG) is present in the chloroplast
RT genome of black pine: a compilation of 32 tRNA genes from black pine
RT chloroplasts.";
RL Curr. Genet. 26:153-158(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95024047; PubMed=7937893;
RA Wakasugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T.,
RA Sugiura M.;
RT "Loss of all ndh genes as determined by sequencing the entire
RT chloroplast genome of the black pine Pinus thunbergii.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).
RN [4]
RP SEQUENCE FROM N.A.
RA Sugiura M.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=94138245; PubMed=8305874;
RA Li J., Goldschmidt-Clermont M., Timko M.;
RT "Chloroplast-encoded chlB is required for light-independent
RT protochlorophyllide reductase activity in Chlamydomonas reinhardtii.";
RL Plant Cell 5:1817-1829(1993).
DR EMBL; D17510; BAA04307.1; -.
DR EMBL; D11467; BAA02023.1; -.
KW Chloroplast.
SQ SEQUENCE 22 AA; 2606 MW; 0A95CB0443BCAEC5 CRC64;

Q37112 Length: 22 December 22, 2002 19:30 Type: P Check: 9119 ..

1 MEYLTTDRSI ECGIYLKKIE SI

!!AA_SEQUENCE 1.0

ID Q957T4 PRELIMINARY; PRT; 22 AA.
AC Q957T4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Ribosomal protein S12 (Fragment).
GN RPS12.
OS Abies alba (Edeltanne) (European silver fir).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=45372;
RN [1]
RP SEQUENCE FROM N.A.
RA Liepelt S.;
RT "Sequence analysis of chosen regions of the mitochondrial genome of
RT Abies alba.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF359456; AAK48935.1; -.
KW Mitochondrion.
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2711 MW; 97BF270D8C4E6558 CRC64;

Q957T4 Length: 22 December 22, 2002 19:30 Type: P Check: 9875 ..

1 MPTSNQSIRH GREKKRRTDR TR

!!AA_SEQUENCE 1.0

ID Q957T0 PRELIMINARY; PRT; 23 AA.
AC Q957T0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Ribosomal protein S12 (Fragment).
GN RPS12.
OS Pinus mugo.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=28528;
RN [1]
RP SEQUENCE FROM N.A.
RA Wachowiak W., Liepelt S., Prus-Glowacki W.;
RT "Sequence analysis of chosen regions of the mitochondrial genome of
RT Pinus mugo and P. sylvestris.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF359576; AAK39117.1; -.
KW Mitochondrion.
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2782 MW; 17A7BF270D8C4E65 CRC64;

Q957T0 Length: 23 December 22, 2002 19:30 Type: P Check: 1370 ..

1 MPTSNQSIRH GREKKRRTDR TRA

!!AA_SEQUENCE 1.0

ID Q957S8 PRELIMINARY; PRT; 23 AA.
AC Q957S8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Ribosomal protein S12 (Fragment).
GN RPS12.
OS Pinus sylvestris (Scots pine).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=3349;
RN [1]
RP SEQUENCE FROM N.A.
RA Wachowiak W., Liepelt S., Prus-Glowacki W.;
RT "Sequence analysis of chosen regions of the mitochondrial genome of
RT Pinus mugo and P. sylvestris.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF359577; AAK39119.1; -.
KW Mitochondrion.
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2782 MW; 17A7BF270D8C4E65 CRC64;

Q957S8 Length: 23 December 22, 2002 19:30 Type: P Check: 1370 ..

1 MPTSNQSI RH GREKKRRTDR TRA

!!AA_SEQUENCE 1.0

ID Q9T2K6 PRELIMINARY; PRT; 20 AA.
AC Q9T2K6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE Acyl-(acyl-carrier-protein) hydrolase 33 kDa polypeptide, AH1
DE (EC 3.1.2.14) (Fragment).
OS Cucurbita moschata (Cushaw squash) (Winter crookneck squash).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
OX NCBI_TaxID=3662;
RN [1]
RP SEQUENCE.
RX MEDLINE=93004473; PubMed=1391766;
RA Imai H., Nishida I., Murata N.;
RT "Acyl-(acyl-carrier protein) hydrolase from squash cotyledons specific
RT to long-chain fatty acids: purification and characterization.";
RL Plant Mol. Biol. 20:199-206(1992).
SQ SEQUENCE 20 AA; 1973 MW; AE5304F9CD73D36E CRC64;

Q9T2K6 Length: 20 December 22, 2002 19:30 Type: P Check: 5922 ..

1 GSSSLADRLX LGSLAXDGFS

!!AA_SEQUENCE 1.0

ID Q9T2H3 PRELIMINARY; PRT; 24 AA.
AC Q9T2H3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Chaperonin 21 (Fragment).
OS Spinacia oleracea (Spinach).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RX MEDLINE=95394976; PubMed=7665625;
RA Ryan M.T., Naylor D.J., Hoogenraad N.J., Hoj P.B.;
RL J. Biol. Chem. 270:22037-22043(1995).
SQ SEQUENCE 24 AA; 2642 MW; B797841E1005A51A CRC64;

Q9T2H3 Length: 24 December 22, 2002 19:30 Type: P Check: 2853 ..

1 ATVVAPKYTS IKPTADRVLI KIKE

!!AA_SEQUENCE 1.0

ID Q37852 PRELIMINARY; PRT; 18 AA.
AC Q37852;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE A protein (Fragment).
 OS Bacteriophage R17.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Leviviridae;
 OC Levivirus.
 OX NCBI_TaxID=12026;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=74125852; PubMed=4361645;
 RA Rensing U.F.E., Coulson A., Schoenmakers J.G.G.;
 RT "A sequence of 54 nucleotides from the A-protein cistron of Coliphage-
 RT R17 RNA."
 RL Eur. J. Biochem. 41:431-438(1974).
 DR EMBL; M24814; AAA32177.1; -.
 FT NON_TER 1 1
 FT NON_TER 18 18
 SQ SEQUENCE 18 AA; 2217 MW; CE3DE9FD8852F11E CRC64;

Q37852 Length: 18 December 22, 2002 19:30 Type: P Check: 3123 ..

1 ALRYLALNED RKFRSKHV

!!AA_SEQUENCE 1.0

ID Q42209 PRELIMINARY; PRT; 28 AA.
 AC Q42209;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Ubiquitin conjugating enzyme UBC10 (Fragment).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA; TISSUE=SEEDLING;
 RA Desprez T., Amselem J., Chiapello H., Caboche M., Hofte H.;
 RL Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; Z29145; CAA82396.1; -.
 DR HSSP; P15731; 1QCQ.
 DR InterPro; IPR000608; UBQ_conjugat.
 DR Pfam; PF00179; UQ_con; 1.
 DR ProDom; PD000461; UBQ_conjugat; 1.
 FT NON_TER 1 1
 SQ SEQUENCE 28 AA; 3321 MW; C4000F6BBD329BAF CRC64;

Q42209 Length: 28 December 22, 2002 19:30 Type: P Check: 1580 ..

1 PEIAHMYKTD RAKYESTARS WTQKYAMG

!!AA_SEQUENCE 1.0

ID Q9FZP4 PRELIMINARY; PRT; 29 AA.
 AC Q9FZP4;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE Genomic DNA, chromosome 5, P1 clone:MVP2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";

RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB025636; BAB11485.1; -.
SQ SEQUENCE 29 AA; 3557 MW; 9B75DC6B293119A2 CRC64;

Q9FZP4 Length: 29 December 22, 2002 19:30 Type: P Check: 3486 ..

1 MDDVDDVFSY LLSKEIDEEN EDREPKYVY

!!AA_SEQUENCE 1.0

ID Q39633 PRELIMINARY; PRT; 23 AA.
AC Q39633;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE Catalase (Fragment).
OS Cucumis sativus (Cucumber).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
OX NCBI_TaxID=3659;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COTYLEDON;
RX MEDLINE=96104306; PubMed=8564304;
RA Toyama T., Teramoto H., Takeba G., Tsuji H.;
RT "Cytokinin induces a rapid decrease in the levels of mRNAs for
RT catalase, 3-hydroxy-3-methylglutaryl CoA reductase, lectin and other
RT unidentified proteins in etiolated cotyledons of cucumber.";
RL Plant Cell Physiol. 36:1349-1359(1995).
DR EMBL; D63385; BAA09701.1; -.
FT NON_TER 1 1
SQ SEQUENCE 23 AA; 2445 MW; DD39DAFD58C3AE65 CRC64;

Q39633 Length: 23 December 22, 2002 19:30 Type: P Check: 1444 ..

1 NGSQADRSVG QKLAPHLNVR PSI

!!AA_SEQUENCE 1.0

ID P82195 PRELIMINARY; PRT; 26 AA.
AC P82195;
DT 01-JUN-2000 (TrEMBLrel. 14, Created)
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Chloroplast 50S ribosomal protein L18 beta (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. ALWARO; TISSUE=LEAF;
RX MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482(2000).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 5S RIBOSOMAL RNA.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 11.5 KDA.
CC -!- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 26 26
SQ SEQUENCE 26 AA; 3184 MW; C5664EC1B2AB7C97 CRC64;

P82195 Length: 26 December 22, 2002 19:30 Type: P Check: 7110 ..

1 KAHTRREDRT ARHVRIRKKV EGTPER

!!AA_SEQUENCE 1.0

ID P82196 PRELIMINARY; PRT; 29 AA.
AC P82196;
DT 01-JUN-2000 (TrEMBLrel. 14, Created)
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Chloroplast 50S ribosomal protein L18 alpha (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. ALWARO; TISSUE=LEAF;
RX MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482(2000).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 5S RIBOSOMAL RNA.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 11.5 KDA.
CC -!- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 29 29
SQ SEQUENCE 29 AA; 3517 MW; 86207D55664EC1B2 CRC64;

P82196 Length: 29 December 22, 2002 19:30 Type: P Check: 4154 ..

1 KAHTRREDRT ARHVRIRKKV EGTPERXXL

!!AA_SEQUENCE 1.0

ID P83089 PRELIMINARY; PRT; 19 AA.
AC P83089;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Thylakoid lumenal 25.3 kDa protein (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE, SUBCELLULAR LOCATION, AND MASS SPECTROMETRY.
RC TISSUE=LEAF;
RX PubMed=11719511;
RA Schubert M., Petersson U.A., Haas B.J., Funk C., Schroeder W.P.,
RA Kieselbach T.;
RT "Proteome map of the chloroplast lumen of Arabidopsis thaliana.";
RL J. Biol. Chem. 277:8354-8365(2002).
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST; WITHIN THE THYLAKOID LUMEN.
CC -!- MASS SPECTROMETRY: MW=25300; METHOD=MALDI.
KW Chloroplast; Thylakoid.
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2093 MW; EC33422F8633F1F9 CRC64;

P83089 Length: 19 December 22, 2002 19:30 Type: P Check: 4307 ..

1 AIANAPLLDT TITDRVFFD

!!AA_SEQUENCE 1.0

ID Q9S8D2 PRELIMINARY; PRT; 29 AA.
AC Q9S8D2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CMETI-B=TRYPSIN inhibitor.
OS Cucumis melo (Muskmelon).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
OX NCBI_TaxID=3656;
RN [1]
RP SEQUENCE.
RX MEDLINE=96015144; PubMed=8537309;
RA Lee C.F., Lin J.Y.;
RT "Amino acid sequences of trypsin inhibitors from the melon Cucumis
RT melo.";
RL J. Biochem. 118:18-22(1995).
DR HSSP; P12071; 2ETI.
DR InterPro; IPR000737; Squash.
DR Pfam; PF00299; squash; 1.
DR PRINTS; PR00293; SQUASHINHBTR.
DR ProDom; PD003401; Squash; 1.
DR SMART; SM00286; PTI; 1.
DR PROSITE; PS00286; SQUASH_INHIBITOR; 1.
SQ SEQUENCE 29 AA; 3196 MW; A4BCFF7AA1AC300E CRC64;

Q9S8D2 Length: 29 December 22, 2002 19:30 Type: P Check: 2605 ..

1 VGCPRILMKC KTRDCLTGC TCKRNGYCG

!!AA_SEQUENCE 1.0

ID Q9S898 PRELIMINARY; PRT; 16 AA.
AC Q9S898;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE 23 kDa heat-induced protein (Fragment).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE.
RX MEDLINE=96351184; PubMed=8742333;
RA Sabehat A., Weiss D., Lurie S.;
RT "The correlation between heat-shock protein accumulation and
RT persistence and chilling tolerance in tomato fruit.";
RL Plant Physiol. 110:531-537(1996).
SQ SEQUENCE 16 AA; 1840 MW; 9B0D3F77BFEE6CE3 CRC64;

Q9S898 Length: 16 December 22, 2002 19:30 Type: P Check: 484 ..

1 NTNTQMTAYD QDDRGX

!!AA_SEQUENCE 1.0

ID Q9QVF2 PRELIMINARY; PRT; 12 AA.
AC Q9QVF2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE TRANSFERRIN=PEPTIDE 21 (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=92165927; PubMed=1791188;
RA Cavanaugh P.G., Nicolson G.L.;
RT "Lung-derived growth factor that stimulates the growth of lung-
RT metastasizing tumor cells: identification as transferrin.";
RL J. Cell. Biochem. 47:261-271(1991).
FT NON_TER 1 1
FT NON_TER 12 12
SQ SEQUENCE 12 AA; 1436 MW; 40AD1DFA420AADD3 CRC64;

Q9QVF2 Length: 12 December 22, 2002 19:30 Type: P Check: 5863 ..

!!AA_SEQUENCE 1.0

ID Q9CTN0 PRELIMINARY; PRT; 18 AA.
 AC Q9CTN0;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE B230218P12Rik protein (Fragment).
 GN B230218P12RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=CORPORA QUADRIGEMINA;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AK021007; BAB32275.1; -.
 DR MGD; MGI:1926130; B230218P12Rik.
 FT NON TER 18 18
 SQ SEQUENCE 18 AA; 2136 MW; 72A0BE772BCA48C8 CRC64;

Q9CTN0 Length: 18 December 22, 2002 19:30 Type: P Check: 2989 ..

1 MPVKKKDTDR ALSLLECY

!!AA_SEQUENCE 1.0

ID Q9CS32 PRELIMINARY; PRT; 27 AA.
 AC Q9CS32;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE 2610020J05Rik protein (Fragment).
 GN 2610020J05RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AK019213; BAB31604.1; -.
 DR MGD; MGI:1914299; 2610020J05Rik.
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 3299 MW; 95C1A07254A2C489 CRC64;

Q9CS32 Length: 27 December 22, 2002 19:30 Type: P Check: 8838 ..

1 MAENTDRNQI EKLLNRVKEL EQEVERL

!!AA_SEQUENCE 1.0

ID Q63990 PRELIMINARY; PRT; 25 AA.
 AC Q63990;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hox3.5 homeobox homolog protein (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR;
 RC TISSUE=BONE MORPHOGENETIC PROTEIN-IMPLANTED SUBCUTANEOUS MUSCLE;
 RX MEDLINE=94271262; PubMed=7911662;
 RA Iimura T., Oida S., Takeda K., Maruoka Y., Sasaki S.;
 RT "Changes in homeobox-containing gene expression during ectopic bone
 RT formation induced by bone morphogenetic protein.";
 RL Biochem. Biophys. Res. Commun. 201:980-987(1994).
 DR EMBL; S71286; AAB31006.1; -.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR ProDom; PD000010; Homeobox; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 3111 MW; 008AA81607BD2203 CRC64;

Q63990 Length: 25 December 22, 2002 19:30 Type: P Check: 5506 ..

1 LFNMYLTRER RLEISKTNL TDRQV

!!AA_SEQUENCE 1.0

ID O88226 PRELIMINARY; PRT; 33 AA.
 AC O88226;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Mszf76 (Fragment).
 GN MSZF76.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL6;
 RX MEDLINE=98296253; PubMed=9630514;
 RA Agata Y., Matsuda E., Shimizu A.;

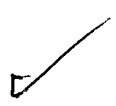
RT "Rapid and efficient cloning of cDNAs encoding Krueppel-like zinc
RT finger proteins by degenerate PCR.";
RL Gene 213:55-64(1998).
DR EMBL; AB010321; BAA31377.1; -.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 1.
DR ProDom; PD000003; Znf_C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW DNA-binding; Metal-binding; Zinc-finger.
FT NON_TER 1 1
FT NON_TER 33 33
SQ SEQUENCE 33 AA; 3730 MW; D8D31E5B5C54C538 CRC64;

O88226 Length: 33 December 22, 2002 19:30 Type: P Check: 2429 ..

1 IHTGEKPYRC AECGKAFTDR SNLIKHQTTH TGE

!!AA_SEQUENCE 1.0

ID Q9JIU0 PRELIMINARY; PRT; 32 AA.
AC Q9JIU0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Protocadherin-T5 (Fragment).
GN PCDH-T5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FISCHER344; TISSUE=TESTIS;
RX MEDLINE=20114371; PubMed=10650949;
RA Johnson K.J., Patel S.R., Boekelheide K.;
RT "Multiple cadherin superfamily members with unique expression profiles
RT are produced in rat testis.";
RL Endocrinology 141:675-683(2000).
DR EMBL; AF177699; AAF87074.1; -.
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 1.
FT NON_TER 1 1
FT NON_TER 32 32
SQ SEQUENCE 32 AA; 3369 MW; A6159FED42844051 CRC64;




Q9JIU0 Length: 32 December 22, 2002 19:30 Type: P Check: 1050 ..

1 ALQAFEFHVG ATDRGSPALS SQALVRVVVL DN

!!AA_SEQUENCE 1.0

ID Q9QX46 PRELIMINARY; PRT; 25 AA.
AC Q9QX46;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Amyloid protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ; TISSUE=LIVER;
RX MEDLINE=20102720; PubMed=10636908;
RA Bergsdorf C., Paliga K., Kreger S., Masters C.L., Beyreuther K.;
RT "Identification of cis-Elements Regulating Exon 15 Splicing of the
RT Amyloid Precursor Protein Pre-mRNA.";
RL J. Biol. Chem. 275:2046-2056(2000).
DR EMBL; AF199005; AAF20194.1; -.
DR EMBL; AF199003; AAF20194.1; JOINED.



DR EMBL; AF199004; AAF20194.1; JOINED.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 2608 MW; 276F026D582421D0 CRC64;

Q9QX46 Length: 25 December 22, 2002 19:30 Type: P Check: 4770 ..

1 ENEVEPVDAR PAADRGLTTR PGSGL

!!AA_SEQUENCE 1.0

ID Q9ET00 PRELIMINARY; PRT; 20 AA.
AC Q9ET00;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE EIF4H (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RA Green E.D.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF289665; AAF99335.1; -.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2188 MW; 8AFFAC0FB953E8F5 CRC64;

Q9ET00 Length: 20 December 22, 2002 19:30 Type: P Check: 5807 ..

1 MADFDTYDDR AYSSFGGGRG

!!AA_SEQUENCE 1.0

ID Q9Z1I5 PRELIMINARY; PRT; 19 AA.
AC Q9Z1I5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE Calcineurin A alpha (EC 3.1.3.16) (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=93111954; PubMed=1335233;
RA Chang C., Takeda T., Mukai H., Shuntoh H., Kuno T., Tanaka C.;
RT "Molecular cloning and characterization of the promoter region of the
RT calcineurin A alpha gene."
RL Biochem. J. 288:801-805(1992).
DR EMBL; D10480; BAA01283.1; -.
KW Hydrolase.
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2115 MW; F905F8B134CCEC57 CRC64;

Q9Z1I5 Length: 19 December 22, 2002 19:30 Type: P Check: 4927 ..

1 MSEPKAIDPK LSTTDVVK

!!AA_SEQUENCE 1.0

ID Q62256 PRELIMINARY; PRT; 18 AA.
AC Q62256;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Spermatogenic-specific proenkephalin.
GN PENK-RS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90287163; PubMed=2355920;
RA Kilpatrick D.L., Zinn S.A., Fitzgerald M., Higuchi H., Sabol S.L.,
RA Meyerhardt J.;
RT "Transcription of the rat and mouse proenkephalin genes is initiated
RT at distinct sites in spermatogenic and somatic cells.";
RL Mol. Cell. Biol. 10:3717-3726(1990).
DR EMBL; M55181; AAA40127.1; -.
DR MGD; MGI:104628; Penk-rs.
SQ SEQUENCE 18 AA; 2043 MW; B96E10CC7049FA76 CRC64;

Q62256 Length: 18 December 22, 2002 19:30 Type: P Check: 3193 ..

1 MSSGKQDSPW EDRIPPGR

!!AA_SEQUENCE 1.0

ID Q63985 PRELIMINARY; PRT; 25 AA.
AC Q63985;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hox1.8 homeobox homolog protein (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94271262; PubMed=7911662;
RA Iimura T., Oida S., Takeda K., Maruoka Y., Sasaki S.;
RT "Changes in homeobox-containing gene expression during ectopic bone
RT formation induced by bone morphogenetic protein.";
RL Biochem. Biophys. Res. Commun. 201:980-987(1994).
DR EMBL; S71275; AAB31001.1; -.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1 1
SQ SEQUENCE 25 AA; 3134 MW; C54E951607BD27B1 CRC64;



Q63985 Length: 25 December 22, 2002 19:30 Type: P Check: 5721 ..

1 LFNMYLTRER RLEISRSVHL TDRQV

!!AA_SEQUENCE 1.0

ID Q61159 PRELIMINARY; PRT; 32 AA.
AC Q61159;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NN8-AG (Fragment).
GN RRG1 OR NN8-AG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV;
RX MEDLINE=96315660; PubMed=8754834;
RA Shago M., Giguere V.;
RT "Isolation of a novel retinoic acid-responsive gene by selection of
RT genomic fragments derived from CpG-island-enriched DNA.";
RL Mol. Cell. Biol. 16:4337-4348(1996).
DR EMBL; U50384; AAB38132.1; -.
DR MGD; MGI:108048; Rrg1.
FT NON_TER 32 32

SQ SEQUENCE 32 AA; 3470 MW; 9571FBD80B1FCA84 CRC64;

Q61159 Length: 32 December 22, 2002 19:30 Type: P Check: 837 ..

1 MAASMCDVFS FCVGVADRAR GSVEVRYVDS IK

!!AA_SEQUENCE 1.0

ID Q923H1 PRELIMINARY; PRT; 40 AA.
AC Q923H1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Aristalless-related homeobox protein Arx (Fragment).
OS Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RA Ohira R.H., Zhang Y.H., Guo W., Dipple K., Shih S., Doerr J.,
RA Huang B.-L., Fu L., Abu-Khalil A., Geschwind D., McCabe E.;
RT "Human ARX gene: genomic characterization and expression."
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY038070; AAK93900.1; -.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1 1
FT NON_TER 40 40
SQ SEQUENCE 40 AA; 3815 MW; 87A852768DB0CB6E CRC64;

Q923H1 Length: 40 December 22, 2002 19:30 Type: P Check: 2032 ..

1 PTPAVEGAVA SGALADPATA AADRRASSIA AXGSRPRSMP

!!AA_SEQUENCE 1.0

ID Q99KX5 PRELIMINARY; PRT; 27 AA.
AC Q99KX5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 2.8 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC003968; AAH03968.1; -.
KW Hypothetical protein.
SQ SEQUENCE 27 AA; 2832 MW; 4008F368AF868B63 CRC64;

Q99KX5 Length: 27 December 22, 2002 19:30 Type: P Check: 9106 ..

1 MASLPVVGVSQ VPADRGYLGP RPAVQEH

!!AA_SEQUENCE 1.0

ID Q9QVD0 PRELIMINARY; PRT; 24 AA.
AC Q9QVD0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Vitronectin (Fragment).
OS Cavia (guinea pigs).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae.
OX NCBI_TaxID=10140;
RN [1]
RP SEQUENCE.
RX MEDLINE=92207982; PubMed=1372829;

RA Nakashima N., Miyazaki K., Ishikawa M., Yatohgo T., Ogawa H.,
RA Uchibori H., Matsumoto I., Seno N., Hayashi M.;
RT "Vitronectin diversity in evolution but uniformity in ligand binding
RT and size of the core polypeptide.";
RL Biochim. Biophys. Acta 1120:1-10(1992).
FT NON_TER 1 1
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2736 MW; 0CD91505B6F22D0A CRC64;

Q9QVD0 Length: 24 December 22, 2002 19:30 Type: P Check: 3385 ..

1 XXESXKGRXT EGFNADRKXQ XXEL

!!AA_SEQUENCE 1.0

ID Q9QV89 PRELIMINARY; PRT; 32 AA.
AC Q9QV89;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE SLP-14=FATTY acid-binding protein (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=93191717; PubMed=8447836;
RA Bansal M.P., Medina D.;
RT "Expression of fatty acid-binding proteins in the developing mouse
RT mammary gland.";
RL Biochem. Biophys. Res. Commun. 191:61-69(1993).
DR HSSP; P05413; 1HMT.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR Pfam; PF00061; lipocalin; 1.
FT NON_TER 1 1
FT NON_TER 32 32
SQ SEQUENCE 32 AA; 3495 MW; 67D23A0146E46AFE CRC64;

Q9QV89 Length: 32 December 22, 2002 19:30 Type: P Check: 581 ..

1 EISFQLGVEF DEVTADDRKV KSVVTLDGK LV

!!AA_SEQUENCE 1.0

ID Q9QV71 PRELIMINARY; PRT; 17 AA.
AC Q9QV71;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Lactase-phlorizin hydrolase (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=93293888; PubMed=8514793;
RA Dudley M.A., Hachey D.L., Quaroni A., Hutchens T.W., Nichols B.L.,
RA Rosenberger J., Perkinson J.S., Cook G., Reeds P.J.;
RT "In vivo sucrase-isomaltase and lactase-phlorizin hydrolase turnover
RT in the fed adult rat.";
RL J. Biol. Chem. 268:13609-13616(1993).
SQ SEQUENCE 17 AA; 1947 MW; 08941FD316F9692F CRC64;

Q9QV71 Length: 17 December 22, 2002 19:30 Type: P Check: 1363 ..

1 DWEDRNFIAA GPLTNDL

!!AA_SEQUENCE 1.0

ID Q61461 PRELIMINARY; PRT; 25 AA.
AC Q61461;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Cytochrome P-450b (Phenobarbital-inducible) (Fragment).
 GN CYP2B13.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84207435; PubMed=6547088;
 RA Stupans I., Ikeda T., Kessler D.J., Nebert D.W.;
 RT "Characterization of a cDNA clone for mouse phenobarbital-inducible
 RT cytochrome P-450b.";
 RL DNA 3:129-137(1984).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; K02409; AAA37510.1; -.
 DR MGD; MGI:88599; Cyp2b13.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; p450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 2979 MW; B9DFF7A57355A1B0 CRC64;

Q61461 Length: 25 ,December 22, 2002 19:30 Type: P Check: 4925 ..

1 SHRLPTLDDR SKMPYTDAMI HEIQR

!!AA_SEQUENCE 1.0

ID Q65747 PRELIMINARY; PRT; 19 AA.
 AC Q65747;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Outer coat protein (VP2) (Fragment).
 OS Bluetongue virus.
 OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
 OX NCBI_TaxID=12591;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88265863; PubMed=2838961;
 RA Gould A.R., Hyatt A.D., Eaton B.T.;
 RT "Morphogenesis of a bluetongue virus variant with an amino acid
 RT alteration at a neutralization site in the outer coat protein, VP2.";
 RL Virology 165:23-32(1988).
 DR EMBL; M21355; AAA42845.1; -.
 DR InterPro; IPR001742; Orbi_VP2.
 DR Pfam; PF00898; Orbi_VP2; 1.
 KW Coat protein.
 FT NON_TER 1 1
 FT NON_TER 19 19
 SQ SEQUENCE 19 AA; 2608 MW; F0428841265D3061 CRC64;

Q65747 Length: 19 December 22, 2002 19:30 Type: P Check: 4484 ..

1 ERLKIFEHRN QRRDEDRFY

!!AA_SEQUENCE 1.0

ID Q66538 PRELIMINARY; PRT; 35 AA.
 AC Q66538;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE 3' proximal protein (Fragment).
 OS Ebola virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae;
 OC Ebola-like viruses.
 OX NCBI_TaxID=11268;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=86124724; PubMed=3946083;
RA Kiley M.P., Wilusz J., McCormick J.B., Keene J.D.;
RT "Conservation of the 3' terminal nucleotide sequences of Ebola and
RT Marburg virus.";
RL Virology 149:251-254(1986).
DR EMBL; M33062; AAA42976.1; -.
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 4142 MW; F7778E5FC6283092 CRC64;

Q66538 Length: 35 December 22, 2002 19:30 Type: P Check: 8137 ..

1 MRKINNFLSL KFDDRNLKLL LICNHTVDS EPHTS

!!AA_SEQUENCE 1.0
ID Q83622 PRELIMINARY; PRT; 9 AA.
AC Q83622;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 1.1 kDa protein (Fragment).
OS Murray valley encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxID=11079;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88118912; PubMed=2828633;
RA Hahn C.S., Hahn Y.S., Rice C.M., Lee E., Dalgarno L., Strauss E.G.,
RA Strauss J.H.;
RT "Conserved elements in the 3' untranslated region of flavivirus RNAs
RT and potential cyclization sequences.";
RL J. Mol. Biol. 198:33-41(1987).
DR EMBL; M35172; AAA66627.1; -.
KW Hypothetical protein.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1055 MW; FF36D40AAB05A2C1 CRC64;

Q83622 Length: 9 December 22, 2002 19:30 Type: P Check: 3517 ..

1 THVSEDRVL

!!AA_SEQUENCE 1.0
ID Q84254 PRELIMINARY; PRT; 22 AA.
AC Q84254;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE X protein (Fragment).
OS Bovine papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10571;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89067912; PubMed=2848926;
RA Stamps A.C., Campo M.S.;
RT "Mapping of two novel transcripts of Bovine papillomavirus type 4.";
RL J. Gen. Virol. 69:3033-3045(1988).
DR EMBL; M35264; AAA46926.1; -.
FT NON_TER 1 1
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2273 MW; 2801BC23480C9CF9 CRC64;

Q84254 Length: 22 December 22, 2002 19:30 Type: P Check: 9066 ..

1 NAGPKPGTTP EDVADRPDL PE

!!AA_SEQUENCE 1.0
ID Q9PWC1 PRELIMINARY; PRT; 39 AA.
AC Q9PWC1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE P55-related MAGUK protein DLG3 (Fragment).
 GN DLG3.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Koenig C., Yan Y.L., Postlethwait J., Wendler S., Campos-Ortega J.A.;
 RT "A recessive mutation leading to vertebral ankylosis is associated
 RT with amino acid exchanges in the zebrafish homologue of the human
 RT membrane associated guanylate kinase protein DLG3."
 RL Mech. Dev. 84:1-12(1999).
 DR EMBL; AF124436; AAD39393.1; -.
 DR InterPro; IPR004172; L27.
 DR Pfam; PF02828; L27; 1.
 FT NON_TER 1 1
 FT NON_TER 39 39
 SQ SEQUENCE 39 AA; 4477 MW; 6A9D5C428F5A2DBF CRC64;

Q9PWC1 Length: 39 December 22, 2002 19:30 Type: P Check: 9645 ..

1 LYELLSVLPS QLQPHVESPD DRSFLHAMFG ERSLSLVK

!!AA_SEQUENCE 1.0

ID O93438 PRELIMINARY; PRT; 29 AA.
 AC O93438;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Lbx1 protein (Fragment).
 GN LBX1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dietrich S., Schubert F.R., Healy C., Sharpe P.T., Lumsden A.;
 RT "Specification of the hypaxial musculature."
 RL Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Schubert F.R., Dietrich S., Chapman S.C., Lumsden A.;
 RT "Expression of the Lbx1 gene in the chick embryo."
 RL Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AJ222831; CAA11016.1; -.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR ProDom; PD000010; Homeobox; 1.
 FT NON_TER 1 1
 SQ SEQUENCE 29 AA; 3352 MW; CABBC02415F1D4F7 CRC64;

O93438 Length: 29 December 22, 2002 19:30 Type: P Check: 3493 ..

1 RFLYQKYLSP ADRDQIAQQL GLTNAQVIT

!!AA_SEQUENCE 1.0

ID Q90297 PRELIMINARY; PRT; 27 AA.
 AC Q90297;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Homeobox (Fragment).
 OS Carassius auratus (Goldfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Carassius.
 OX NCBI_TaxID=7957;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=RETINA;
 RX MEDLINE=93219357; PubMed=8096640;
 RA Levine E.M., Schechter N.;
 RT "Homeobox genes are expressed in the retina and brain of adult
 RT goldfish.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2729-2733(1993).
 DR EMBL; L09693; AAA49178.1; -.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 3338 MW; 55318651A81607BD CRC64;

Q90297 Length: 27 December 22, 2002 19:30 Type: P Check: 9410 ..

1 LFNMYLTRER RLEISKSINL TDRQVKI

!!AA_SEQUENCE 1.0
 ID Q90298 PRELIMINARY; PRT; 27 AA.
 AC Q90298;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Homeobox (Fragment).
 OS Carassius auratus (Goldfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Carassius.
 OX NCBI_TaxID=7957;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=RETINA;
 RX MEDLINE=93219357; PubMed=8096640;
 RA Levine E.M., Schechter N.;
 RT "Homeobox genes are expressed in the retina and brain of adult
 RT goldfish.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2729-2733(1993).
 DR EMBL; L09694; AAA49179.1; -.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 3375 MW; 52B7254E951607BD CRC64;

Q90298 Length: 27 December 22, 2002 19:30 Type: P Check: 9642 ..

1 LFNMYLTRER RLEISRSVHL TDRQVKI

!!AA_SEQUENCE 1.0
 ID Q07145 PRELIMINARY; PRT; 27 AA.
 AC Q07145;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Homeodomain (Fragment).
 GN HOMEBOX.
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93317669; PubMed=8101001;

RA Pendleton J.W., Nagai B.K., Murtha M.T., Ruddle F.H.;
 RT "Expansion of the Hox gene family and the evolution of chordates."
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6300-6304(1993).
 DR EMBL; L14895; AAA02530.1; -.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 3322 MW; 5529340A951607BD CRC64;

Q07145 Length: 27 December 22, 2002 19:30 Type: P Check: 9552 ..

1 LFNMYLTRER RLEISRGVNL TDRQVKI

!!AA_SEQUENCE 1.0
 ID Q07150 PRELIMINARY; PRT; 27 AA.
 AC Q07150;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Homeodomain (Fragment).
 GN HOMEBOX.
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93317669; PubMed=8101001;
 RA Pendleton J.W., Nagai B.K., Murtha M.T., Ruddle F.H.;
 RT "Expansion of the Hox gene family and the evolution of chordates."
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6300-6304(1993).
 DR EMBL; L14900; AAA02535.1; -.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 3319 MW; 53CD02F1671607BD CRC64;

Q07150 Length: 27 December 22, 2002 19:30 Type: P Check: 9407 ..

1 LFSMYLTRER RLEISHLLSL TDRQVKI

!!AA_SEQUENCE 1.0
 ID O73591 PRELIMINARY; PRT; 14 AA.
 AC O73591;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hox C10 (Fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WHITE LEGHORN;
 RX MEDLINE=98141813; PubMed=9473273;
 RA Peale F.V. Jr., Mason K., Hunter A.W., Bothwell M.;
 RT "Multiplex display polymerase chain reaction amplifies and resolves
 RT related sequences sharing a single moderately conserved domain."
 RL Anal. Biochem. 256:158-168(1998).
 DR EMBL; U34614; AAC36452.1; -.
 FT NON_TER 1 1
 FT NON_TER 14 14
 SQ SEQUENCE 14 AA; 1673 MW; 81258FC9E81FDA45 CRC64;

1 RLEISKSINL TDRQ

```
!!AA_SEQUENCE 1.0
ID   P83009      PRELIMINARY;          PRT;    13 AA.
AC   P83009;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Phospholemman (FXYP domain-containing ion transport regulator 1)
DE   (Fragment).
OS   Lamna nasus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Lamniformes; Alopiidae;
OC   Lamna.
OX   NCBI_TaxID=7849;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=RECTAL GLAND;
RA   Schuurmans Stekhoven F.M.A.H., Flik G., Wendelaar Bonga S.E.;
RT   "Phospholemman in the rectal gland of sharks.";
RL   Submitted (JUN-2001) to the SWISS-PROT data bank.
CC   -!- FUNCTION: INDUCES A HYPERPOLARIZATION-ACTIVATED CHLORIDE CURRENT
CC       WHEN EXPRESSED IN XENOPUS OOCYTES. MAY HAVE A FUNCTIONAL ROLE IN
CC       MUSCLE CONTRACTION.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MICROSOMAL
CC       MEMBRANE.
CC   -!- PTM: MAJOR PLASMA MEMBRANE SUBSTRATE FOR CAMP-DEPENDENT PROTEIN
CC       KINASE (PK-A) AND PROTEIN KINASE C (PK-C) IN SEVERAL DIFFERENT
CC       TISSUES. PHOSPHORYLATED IN RESPONSE TO INSULIN AND ADRENERGIC
CC       STIMULATION (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE FXYP FAMILY.
DR   InterPro; IPR000272; ATP1G1_PLM_MAT8.
DR   PROSITE; PS01310; FXYP; PARTIAL.
KW   Transmembrane; Phosphorylation; Ionic channel; Ion transport;
KW   Microsome.
FT   NON_TER      13      13
SQ   SEQUENCE    13 AA;  1542 MW;  081373C69724A444 CRC64;
```

P83009 Length: 13 December 22, 2002 19:30 Type: P Check: 7125 ..

1 VSDVPNDDR FTY

```
!!AA_SEQUENCE 1.0
ID   P83010      PRELIMINARY;          PRT;    13 AA.
AC   P83010;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Phospholemman (FXYP domain-containing ion transport regulator 1)
DE   (Fragment).
OS   Triakis scyllium (Leopard shark) (Triakis scyllia).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes; Triakidae;
OC   Triakis.
OX   NCBI_TaxID=30494;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=RECTAL GLAND;
RA   Schuurmans Stekhoven F.M.A.H., Flik G., Wendelaar Bonga S.E.;
RT   "Phospholemman in the rectal gland of sharks.";
RL   Submitted (JUN-2001) to the SWISS-PROT data bank.
CC   -!- FUNCTION: INDUCES A HYPERPOLARIZATION-ACTIVATED CHLORIDE CURRENT
CC       WHEN EXPRESSED IN XENOPUS OOCYTES. MAY HAVE A FUNCTIONAL ROLE IN
CC       MUSCLE CONTRACTION.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MICROSOMAL
CC       MEMBRANE.
CC   -!- PTM: MAJOR PLASMA MEMBRANE SUBSTRATE FOR CAMP-DEPENDENT PROTEIN
CC       KINASE (PK-A) AND PROTEIN KINASE C (PK-C) IN SEVERAL DIFFERENT
CC       TISSUES. PHOSPHORYLATED IN RESPONSE TO INSULIN AND ADRENERGIC
```


CC STIMULATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE FXYD FAMILY.
DR InterPro; IPR000272; ATP1G1_PLM_MAT8.
DR PROSITE; PS01310; FXYD; PARTIAL.
KW Transmembrane; Phosphorylation; Ionic channel; Ion transport;
KW Microsome.
FT NON_TER 13 13
SQ SEQUENCE 13 AA; 1497 MW; 1D140C92C1AE1444 CRC64;

P83010 Length: 13 December 22, 2002 19:30 Type: P Check: 6920 ..

1 AGEPANNE^{DR} FNY

!!AA_SEQUENCE 1.0

ID Q90ZG3 PRELIMINARY; PRT; 27 AA.
AC Q90ZG3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE HoxC11a (Fragment).
GN MF-HOXC11A.
OS Oryzias latipes (Medaka fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kurosawa G., Yamada K., Ishiguro H., Hori H.;
RT "isolation of BAC Clones That Cover Seven Hox Clusters in Medaka
RT Genome.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB055740; BAB62853.1; -.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 3368 MW; 548C73EA983E9158 CRC64;

Q90ZG3 Length: 27 December 22, 2002 19:30 Type: P Check: 9516 ..

1 FFN^{VY}INKEK RLQLSRMLNL TDRQVKI

!!AA_SEQUENCE 1.0

ID Q9PS38 PRELIMINARY; PRT; 20 AA.
AC Q9PS38;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE PO GLYCOPROTEIN=MYELIN major structural protein (Fragment).
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE.
RX MEDLINE=93002191; PubMed=1382532;
RA Karthigasan J., Bauer T.K., Teplow D.B., Saavedra R.A.,
RA Kirschner D.A.;
RT "Phylogenetically conserved amino acids of MBP and P0 from amphibian
RT myelin.";
RL J. Mol. Neurosci. 3:185-188(1992).
FT NON_TER 1 1
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2351 MW; AB238ACED7837676 CRC64;

Q9PS38 Length: 20 December 22, 2002 19:30 Type: P Check: 6634 ..

1 IEVYTDREIQ SNVGSKVHLY

!!AA_SEQUENCE 1.0
ID Q9PS32 PRELIMINARY; PRT; 23 AA.
AC Q9PS32;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Nonmuscle alpha-actinin 115 kDa isoform (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE=93100311; PubMed=1334489;
RA Imamura M., Masaki T.;
RT "A novel nonmuscle alpha-actinin. Purification and characterization of
RT chicken lung alpha-actinin."
RL J. Biol. Chem. 267:25927-25933(1992).
FT NON_TER 1 1
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2823 MW; EBED28437F668363 CRC64;

Q9PS32 Length: 23 December 22, 2002 19:30 Type: P Check: 1521 ..

1 LASDLLEWIR RTIPWLED RS PQK

!!AA_SEQUENCE 1.0
ID Q9PRV3 PRELIMINARY; PRT; 24 AA.
AC Q9PRV3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE NAP-185 NEUROGLIA-associated protein (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE=95050502; PubMed=7961672;
RA Solca F.F., Lurie D.I., Diltz C.D., Johnson R.S., Kumar S.,
RA Rubel E.W., Fischer E.H.;
RT "Identification and purification of a chicken brain neuroglia-
RT associated protein."
RL J. Biol. Chem. 269:27559-27565(1994).
SQ SEQUENCE 24 AA; 2404 MW; 73E0E6067F253442 CRC64;

Q9PRV3 Length: 24 December 22, 2002 19:30 Type: P Check: 2489 ..

1 DGGEDRDAAV EEAVLGTGGC RTPK

!!AA_SEQUENCE 1.0
ID Q9DU42 PRELIMINARY; PRT; 25 AA.
AC Q9DU42;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Rev protein (Fragment).
GN REV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirus; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MIDU2;
RX MEDLINE=21002573; PubMed=11118069;
RA Motomura K., Kusagawa S., Kato K., Nohtomi K., Lwin H.H., Tun K.M.,
RA Thwe M., Oo K.Y., Lwin S., Kyaw O., Zaw M., Nagai Y., Takebe Y.;
RT "Emergence of new forms of human immunodeficiency virus type 1

RT intersubtype recombinants in central myanmar.";
RL AIDS Res. Hum. Retroviruses 16:1831-1843(2000).
DR EMBL; AB043899; BAB19213.1; -.
DR InterPro; IPR000625; REV_protein.
DR Pfam; PF00424; REV; 1.
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 2964 MW; 0F92F42A1DDF2758 CRC64;

Q9DU42 Length: 25 December 22, 2002 19:30 Type: P Check: 5339 ..

1 MAGRSED RDE ELLKTVRLIK LLYQS

!!AA_SEQUENCE 1.0

ID Q70140 PRELIMINARY; PRT; 9 AA.
AC Q70140;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Tat protein (Fragment).
GN TAT.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=020;
RX MEDLINE=95194694; PubMed=7888189;
RA Gao F., Yue L., Craig S., Thornton C.L., Robertson D.L.,
RA McCutchan F.E., Bradac J.A., Sharp P.M., Hahn B.H.;
RT "Genetic variation of HIV type 1 in four World Health Organization-
RT sponsored vaccine evaluation sites: generation of functional envelope
RT (glycoprotein 160) clones representative of sequence subtypes A, B, C,
RT and E. WHO Network for HIV Isolation and Characterization.";
RL AIDS Res. Hum. Retroviruses 10:1359-1368(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=020;
RX MEDLINE=96190564; PubMed=8627686;
RA Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S.,
RA Karlsson G., Sodroski J., Morgado M., Galvao-Castro B.,
RA von Briesen H., Beddows S., Weber J., Sharp P.M., Shaw G.M.,
RA Hahn B.H.;
RT "Molecular cloning and analysis of functional envelope genes from
RT human immunodeficiency virus type 1 sequence subtypes A through G. The
RT WHO and NIAID Networks for HIV Isolation and Characterization.";
RL J. Virol. 70:1651-1657(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=020;
RA Allen E.E.;
RL Submitted (APR-1994) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U08794; AAB05175.1; -.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1098 MW; 5B76D40AB1AB01A3 CRC64;

Q70140 Length: 9 December 22, 2002 19:30 Type: P Check: 3335 ..

1 SKTETDRFD

!!AA_SEQUENCE 1.0

ID O25575 PRELIMINARY; PRT; 23 AA.
AC O25575;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein HP0917.
GN HP0917.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=26695 / ATCC 700392;
 RX MEDLINE=97394467; PubMed=9252185;
 RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
 RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen *Helicobacter*
 RT *pylori*.";
 RL Nature 388:539-547(1997).
 DR EMBL; AE000601; AAD07968.1; -.
 DR TIGR; HP0917; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 23 AA; 2717 MW; FC3042379DA2FD44 CRC64;

025575 Length: 23 December 22, 2002 19:30 Type: P Check: 575 ..

1 MSPLTPLRNP LTQEDRFFQE IIA

!!AA_SEQUENCE 1.0
 ID O06283 PRELIMINARY; PRT; 27 AA.
 AC O06283;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein Rv3599c.
 GN RV3599C OR MTCY07H7B.23.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaiia F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of *Mycobacterium tuberculosis* from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 DR EMBL; Z95557; CAB08945.1; -.
 DR TubercuList; Rv3599c; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 27 AA; 2894 MW; 1746F2AA97B95C6B CRC64;

O06283 Length: 27 December 22, 2002 19:30 Type: P Check: 8577 ..

1 MPASSLGTGS PAADRLDATH ERRREVI

!!AA_SEQUENCE 1.0
 ID Q9KTZ5 PRELIMINARY; PRT; 33 AA.
 AC Q9KTZ5;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical protein VC0735.
 GN VC0735.
 OS *Vibrio cholerae*.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.

OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RT *cholerae*";
 RL Nature 406:477-483(2000).
 DR EMBL; AE004159; AAF93900.1; -.
 DR TIGR; VC0735; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 33 AA; 3661 MW; A3AD77F43D6FE40C CRC64;

Q9KTZ5 Length: 33 December 22, 2002 19:30 Type: P Check: 3787 ..

1 MPVGSSIPCS ISTALADRKV VSRWCFRTL LVL

!!AA_SEQUENCE 1.0

ID Q9KLI3 PRELIMINARY; PRT; 32 AA.
 AC Q9KLI3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical protein VCA0761.
 GN VCA0761.
 OS *Vibrio cholerae*.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RT *cholerae*";
 RL Nature 406:477-483(2000).
 DR EMBL; AE004404; AAF96659.1; -.
 DR TIGR; VCA0761; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 32 AA; 3898 MW; E08F60B7C8C67F83 CRC64;

Q9KLI3 Length: 32 December 22, 2002 19:30 Type: P Check: 1300 ..

1 MPDRLLRFSA ICSTDRRKQT NLFSSDLKQQ WR

!!AA_SEQUENCE 1.0

ID Q8X407 PRELIMINARY; PRT; 13 AA.
 AC Q8X407;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein z0899.
 GN z0899.
 OS *Escherichia coli* O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC *Escherichia*.
 OX NCBI_TaxID=83334;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).
 DR EMBL; AE005252; AAG55068.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 13 AA; 1575 MW; 5127D365A6BBD763 CRC64;

Q8X407 Length: 13 December 22, 2002 19:30 Type: P Check: 6940 ..

1 MSTDRKPVML LFH

!!AA_SEQUENCE 1.0
 ID Q8VJ09 PRELIMINARY; PRT; 35 AA.
 AC Q8VJ09;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein MT3536.
 GN MT3536.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / OSHKOSH;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AE007158; AAK47876.1; -.
 DR TIGR; MT3536; -.
 KW Hypothetical protein.
 SQ SEQUENCE 35 AA; 4204 MW; 897851C24908FC03 CRC64;

Q8VJ09 Length: 35 December 22, 2002 19:30 Type: P Check: 8356 ..

1 MKYKLAILDE YDRADRTERG AILRRENLYS SLLTE

!!AA_SEQUENCE 1.0
 ID Q9HNL3 PRELIMINARY; PRT; 32 AA.
 AC Q9HNL3;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE Vng2049c.
 GN VNG2049C.
 OS Halobacterium sp. (strain NRC-1).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20504483; PubMed=11016950;
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,

RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005097; AAG20207.1; -.
KW Complete proteome.
SQ SEQUENCE 32 AA; 3719 MW; D0F9677E650A578A CRC64;

Q9HNL3 Length: 32 December 22, 2002 19:30 Type: P Check: 9872 ..

1 MVRVPVTDAG RAEMRERADR LETTAAFWRL VD

